VOL. XXI APRIL, 1937 No. 2

THE JOURNAL OF APPLIED PSYCHOLOGY

JAMES P. PORTER, *Editor*
OHIO UNIVERSITY ATHENS, OHIO

With a Board of Cooperating Editors
and
PAUL S. ACHILLES HENRY C. LINK
Associate Editors Representing
THE PSYCHOLOGICAL CORPORATION

SOME INTERESTING CONTRIBUTIONS IN THIS ISSUE:

ROWENA RIPIN AND PAUL F. LAZARSFELD. The Tactile-Kinaesthetic Perception of Fabrics with Emphasis on Their Relative Pleasantness

A. R. LAUER. Fact and Fancy Regarding Driver Testing Procedures

HERBERT A. TOOPS AND S. EDSON HAVEN. Viewing the Traffic Problem

DOROTHY M. ANDREW. An Analysis of the Minnesota Vocational Test for Clerical Workers, II

PUBLISHED BI-MONTHLY BY THE EDITOR
OHIO UNIVERSITY, ATHENS, OHIO
PRINTED BY THE SCIENTIFIC PRESS PRINTING COMPANY, LANCASTER, PA.

Entered as second-class matter at the post office at Athens, Ohio, under the act of March 3, 1879. Additional entry May 7, 1936, at the post office at Lancaster, Pa. Copyright 1936 by James P. Porter.

The Journal of Applied Psychology

COOPERATING EDITORS

W. F. Book, Indiana University, W. V. Bingham, Personnel Research Federation, New York; M. E. Broom, State Teachers College, San Diego, Calif.; H. A. Bursy, Ohio State University; Norman R. Gault, Northwestern University; William Healy, Judge Foundation, Juvenile Court, Boston; H. L. Hollingworth, Columbia University; J. E. W. Wallin, Antioch University; Harold S. Langfeld, Princeton University; James H. Leuba, Bryn Mawr College; Henry T. McComas, Johns Hopkins University; Walter D. Scott, Northwestern University; C. E. Seashore, University of Iowa; Daniel Starch, New York; Edward K. Strong, Jr., Leland Stanford Jr. University; L. M. Terman, Leland Stanford Jr. University; E. L. Thorndike, Columbia University; Joseph Jastrow, New School of Social Research, New York; G. M. Whipple, Denver, Mass.; Helen Thompson Woolley, Teachers College, Columbia University.

The scope of THE JOURNAL OF APPLIED PSYCHOLOGY will include the following: (a) The application of psychology to vocational activities, such as law, art, public speaking, industrial and commercial work, and problems of business efficiency. (b) Studies of individual personalities, such as types of character, specific talents, genius, and individual differences, including the problems of mental diagnosis and vocational prognosis. (c) The influence of general environmental conditions, such as climate, weather, humidity, temperature; also such conditions as nutrition, fatigue, etc. (d) The psychology of everyday activities, such as reading, writing, speaking, singing, playing games or playing musical instruments, sewing, etc.

Communications, manuscripts, reviews, abstracts, letters, books and offprints for review should be addressed to Dr. James P. Porter, Ohio University, Athens, Ohio. Contributors are requested to follow the conventional method of citation employed by this JOURNAL; in every case the year, volume and pages should be specified.

THE JOURNAL OF APPLIED PSYCHOLOGY is published bi-monthly. The subscription price is $4.00 a year. Single numbers $1.25 to $2.00, depending upon the number of pages. Remittances and business communications should be addressed to the publisher, James P. Porter, Ohio University, Athens, Ohio.

Vol. XXI APRIL, 1937 No. 2

THE JOURNAL OF APPLIED PSYCHOLOGY

Edited by
JAMES P. PORTER, Ohio University
Athens, Ohio

TABLE OF CONTENTS

DOROTHY M. ANDREW. An Analysis of the Minnesota Vocational Test for Clerical Workers II 139

A. R. LAUER. Fact and Fancy Regarding Driver Testing Procedures 173

HERBERT A. TOOPS AND S. EDSON HAVEN. Viewing the Traffic Problem 185

ROWENA RIPIN AND PAUL F. LAZARSFELD. The Tactile-Kinaesthetic Perception of Fabrics with Emphasis on Their Relative Pleasantness 198

HERBERT A. TOOPS. The Factor of Mechanical Arrangement and Typography in Questionnaires 225

HERMAN A. COPELAND. Validating Two Tests for Census Enumeration 230

RICHARD A. TRIPLETT. Do You Know the Multiplication Table? 233

News and Notes 241

Book Reviews 245

New Books and Pamphlets 250

AN ANALYSIS OF THE MINNESOTA VOCATIONAL TEST FOR CLERICAL WORKERS—II

DOROTHY M. ANDREW
Pennsylvania College for Women

EARLIER results have indicated that the Minnesota Vocational Test for Clerical Workers is a reliable and valid measure of clerical aptitude, and that it is relatively free from the influence of clerical training and experience (3). Both the clerical name and number checking tests are relatively independent of academic ability when homogeneous groups are considered. For more heterogeneous groups, only clerical number checking is independent of academic ability. The clerical test differentiates clerical workers from workers in the general population and from workers in various non-clerical occupations.

This article reports the second part of a comprehensive study on the nature of clerical ability.* Four problems have been singled out for study, namely:

1, to discover what aptitudes the clerical test is measuring and what traits are closely related to it
2, to determine whether the clerical test is measuring a unique trait which is relatively independent of other tested abilities
3, to determine the value of multiple factor and tetrad analysis techniques in the study of the overlapping among abilities and of the number of factors involved

* This is the second article taken from the author's Ph.D. dissertation on file in the office of the Graduate School of the University of Minnesota. The first part appears in the February 1937 issue of this JOURNAL.

in a test battery which measures academic, clerical, dexterity, and mechanical abilities

4, to evaluate the contributions and the limitations of factor analysis techniques in a vocational testing program.

For these studies a group of young women clerical workers, homogeneous as to sex, age, and general occupation has served as the experimental group. All tests used were tested as to their reliability and capability of giving a normal distribution of scores for this group before inclusion in this experiment.

To study the first problem, namely, that of the nature of the test and of the abilities to which it is related, tests were selected which measured relatively simple abilities and were revised to meet requirements of reliability. These tests included number and letter cancellation, letter and number substitution, arithmetic, spelling, number group checking, and addition. Four feature discrimination tests from the Michigan Non-Verbal Series devised by E. B. Greene (9) were also used because they supposedly measured an ability to observe and compare, which ability might presumably be related to the clerical test.

The uniqueness of the clerical test was studied by observing its relation to the Pressey Education Tests, the O'Connor Finger and Tweezer Dexterity Tests, and the Minnesota Spatial Relations Test, all of which tests were used by the Minnesota Employment Stabilization Research Institute in its testing program (8).

These two problems were answered in Part I of this investigation. There it was indicated that the average intercorrelation for all of the variables was + .27. The uncorrected correlations ranged from about .00 to + .84; the highest correlation being between number and letter cancellation.

Correlational analysis indicated the possibility that the test battery is measuring at least four relatively independent factors, those of clerical, academic, spatial, and dexterity abilities.

There also is the probability that the clerical number and name checking tests are measuring two factors, one a numerical ability and the other a verbal ability. There are indications also that the two dexterity tests, though related are measuring relatively specific abilities. Although the spatial relations test is more closely related to finger and tweezer dexterities than are the feature discrimination tests, the feature discrimination tests are more closely associated than spatial relations to clerical and academic abilities. The feature discrimination tests seem to be measuring abilities which are important in academic, clerical, and spatial tests.

The correlations indicate that both clerical experience and education are independent of the abilities measured by the various tests. Academic ability is significantly related to the verbal-clerical and feature discrimination tests but is independent of the dexterities and of the numerical clerical tests. The dexterity tests are somewhat related to spatial ability but are independent of other test groups. Spatial ability is definitely unrelated to the clerical test groups. The feature discrimination tests are related to the clerical and spatial tests, but are independent of the dexterity tests. When the clerical tests are divided into two groups, those dealing with letters and those dealing with numbers, no differences are present between the correlations of the verbal with the verbal and the verbal with the numerical tests.

The third and fourth problems, one being the overlapping among abilities and the number of factors involved in a test battery, which measures academic, clerical, dexterity, and mechanical abilities, and the other being an evaluation of factor analysis techniques in a vocational testing program, are the subject of this paper.

The intercorrelations for this analysis will be studied by the Thurstone "Simplified Multiple Factor Method" (23) and by Spearman's technique of tetrad analysis (10). Factor analysis is considered in this investigation because of its possible con-

tribution to the understanding of the measured abilities and not for any theoretical considerations aimed to disprove or substantiate the techniques or the various theories of factor analysis.

An important question then, is that of the value of factor analysis for the applied psychologist in improving his understanding of testing programs and of their results. An evaluation of the usefulness and limitations of factor analysis techniques as applied to practical problems will be considered. To test the general applicability of the technique of multiple factor analysis, modifications in the data will be made to show how the addition or omission of certain variables will affect the size and number of factors; and an attempt will be made to find out how factors are dependent upon the type of group examined. A comparison of results obtained by tetrad and multiple factor analysis techniques will be made.

Results secured by applying factor analysis to ability tests should be of great importance to occupational testing programs. Tests shown to lack specificity and to overlap abilities measured by other tests may be omitted. In the hope of securing measures of specific abilities, the examiner may substitute other tests which are not included in the battery because of the limitations of testing time. Factor analysis will aid in pointing out which aptitudes are unique and which tests lack specificity. The aim of every comprehensive testing program is to measure adequately all of the important traits of the individual with a minimum expenditure of time. In a vocational testing program, factor analysis techniques will contribute by aiding in the selection of tests measuring specific abilities and in the construction of new tests which will measure one ability or factor, a factor which is independent of other measured abilities.

If individual differences can be resolved into a few broad traits, vocational guidance will be less complex and very different from the guidance techniques required if behavior is found to be made up of numerous specific abilities.

APPLICATIONS OF FACTOR ANALYSIS

Many investigators have applied factor analysis techniques to the intercorrelations of tested abilities either because of interest in practical applications for the analysis of mental traits, or because of interest in the confirmation, challenging, or study of some theory of mental ability, whether it be a two factor or multiple factor theory. A complete survey of the various studies is not feasible, but some of the more important studies forming a background for the present research will be discussed.

In addition to a general factor of intelligence, Spearman has postulated universal abilities of inertia, perseveration, general oscillation, and character force (19). His analyses have indicated the presence of a dozen group factors. The group factors include logical, mechanical, arithmetical, social, and geometrical abilities as well as reaction time, visual-auditory imagery, general memorization, sensory memory, verbal memory, nonverbal symbolic memory, and musical appreciation.

Burt (5) also, believes there are group factors in arithmetical, manual, verbal, linguistic, literary, artistic, and musical abilities.

Kelley (12) agrees with Spearman in believing in the existence of a spatial, numerical, memory, and general factor. The difference between Kelley's and Spearman's position relative to a general factor is one of interpretation, Spearman believing the general factor is due to general nervous energy, and Kelley ascribing it to heterogeneity of the group in race, sex, or age. Spearman and Kelley agree that a large number of specific motor and sensory factors exist.

Wilson (25) found conclusive evidence of a group factor in arithmetical tests, and a suggestion of group factors in verbal and memory tests.

Brown and Stephenson (4) devised mental tests to answer Karl Pearson's idea of testing the factor theory by having a battery of tests measuring twenty non-overlapping abilities.

The verbal tests include measures of paragraph comprehension and inferences and the non-verbal tests include measures of cancellation, arithmetical equations, mutilated pictures, fitting shapes, code or substitution tests. The verbal, unlike the non-verbal tests, did not satisfy the tetrad criterion. The investigators believe, therefore, that the criterion is satisfied only when "past achievement and therefore, the rôle of retentivity and reproduction is rendered as simple as possible as when primarily perceptual tests are used." Non-overlapping tests are most likely to be the perceptual measures.

The relation between verbal and numerical abilities has been studied by Schneck (18). His verbal tests included measures of vocabulary, opposites, analogies, sentence completion, and disarranged sentences. Numerical tests included measures of arithmetical reasoning, number series, number completions, mental multiplication, and equation relations. The average uncorrected intercorrelation for the verbal test was + .49, for numerical tests, + .34, and between verbal and numerical tests, + .14. Schneck found a central factor plus specific factors for the verbal and numerical tests. He holds that verbal and numerical abilities are relatively independent of each other and of memory. The vocabulary test was the best test of verbal ability and the arithmetic test the best measure of numerical ability.

Stephenson reports three studies on the relationships among abilities, one on non-verbal tests, another on verbal tests, and a third on the relation between verbal and non-verbal tests. The tetrad criterion was not satisfied either for the verbal or the non-verbal batteries of tests. Stephenson reports a correlation of + .82 between the verbal and non-verbal tests. The evidence was against any group factor in the non-verbal tests but was definitely in favor of a group factor among the verbal tests. (20) (21) (22)

A study of verbal, numerical, and spatial abilities is reported by Schiller (17). The test battery consisted of four verbal, three numerical, and five spatial-manipulation tests. The

average intercorrelation was +.558. The verbal tests had the highest self-correlations, numerical tests the second, and the spatial tests the lowest intercorrelations. Schiller states that a single common function seems to run through all of the tests. The author reports the presence of verbal, numerical, and spatial factors.

A study of memory tests in relation to verbal, numerical, and spatial abilities has been made by Anastasi (1). The average intercorrelation of the memory tests was +.40. Memory was found to be independent of verbal, numerical, and spatial abilities. Evidence for a central factor in the memory tests was found in the size of the intercorrelations, the application of the tetrad criterion, and the correlations of each test with a central factor. In a later study by Anastasi (2) on the memory factor the results corroborate her former findings of the existence of memory, verbal, and numerical factors and their mutual independence.

Paterson, Elliott, *et al.* (15) have reported a comprehensive study of the measurement of mechanical ability. They constructed and standardized three valid and reliable tests of mechanical ability, a paper form board test, a spatial relations test, and a mechanical assembly test. Paterson and Elliott conclude that the evidence is in favor of specificity of mechanical ability and that it indicates the presence of group factors. Carter (6) in a reply to Spearman states that the Minnesota Mechanical Abilities Investigation shows overlapping group factors.

Jorgenson (11) reports a study in which the Army Alpha, Otis Test for Business Establishments, George Washington Teaching Aptitude Test, the Thurstone Clerical, Thurstone Technical Information, and the McQuarrie Mechanical Ability Tests were given.The correlation between the clerical test and the Army Alpha was +.30 and between the mechanical test and the Army Alpha was +.24. Clerical ability is relatively independent of mechanical ability, as shown by the correlation of +.12. Application of the tetrad difference criterion indi-

cated the absence of both a central factor of memory or clerical ability. Jorgenson states that even in the general intelligence tests there is more that is specific than is general. He cautions against the careless use of tetrad analysis because he finds that the same tests for different groups give varying results.

Thurstone (24) has applied multiple factor analysis to the unpublished data of W. P. Alexander on psychological tests. The Stanford Binet, Pintner-Paterson Scale, Healy Puzzle II, Porteus Maze, Thorndike Reading, Otis Group, Otis Self-Administering, Terman Group, and the Alexander Performance Tests were given to 71 subjects. Two factors, one a verbal, the other a manipulative factor, explain the relationships among the variables. The verbal category includes all of the tests with the exception of the Pintner-Paterson, Healy Puzzle II, Porteus Maze, and the Alexander Performance Tests which measure manipulative abilities.

Thurstone (24) also has applied his multiple factor techniques to Brown's data consisting of twenty tests. Stephenson calculated 15,000 tetrads for these twenty tests and concluded that one of the factors is Spearman's 'G' and the second is verbal ability. Thurstone believes that one of the factors is berbal and the other factor one of visual form perception.

The results of applying factor analysis techniques to the intercorrelations from Brigham's report to the College Entrance Board are also reported by Thurstone (23). Nine tests were used, namely, definitions, arithmetical problems, artificial language, classification, antonyms number series completion, analogies, logical inference, and paragraph reading. The nine tests were reducible to three factors, two of which were a verbal and a numerical factor.

Two studies, one by Perry (16) and the other by Line, Rogers, and Kaplan (13) have been made on relationship between tetrad analysis and Thurstone's multiple factor analysis and on the effect of alteration of variables and corresponding factor loadings. Perry in studying personality, intelligence, and achievement tests concludes that addition of intelligence

and achievement tests to the adjustment questionnaires did not affect the weights of the two main factors. The study yielded evidence for the validity of the Thurstone Technique since it and tetrad, and correlational analyses yielded the same algebraic pattern. Line, Rogers, and Kaplan state that there is "perfect harmony" between tetrad and multiple factor analysis since they put into their table four factors by the Spearman method and isolated four factors by Thurstone's techniques of analysis. Tetrad analysis indicated group factors which could be predicted from factorial loadings. When the experimenters added the National Intelligence Test to the original variables of achievement in school subjects, Pintner Intelligence Test, and tests of perseveration, alertness, speed of execution, and fluency, and recalculated the factors, the results showed again the presence of the four factors with only minor evidence of disturbances.

A review of the literature indicates that fundamental criticisms of factor analysis studies have been concerned with faulty experimental procedure. In the present investigation, an attempt has been made to avoid procedures which might deserve any one of the frequently made criticisms: too small a sampling, use of heterogeneous samples, in age, sex, race, or scholastic training, use of measures which do not yield objective scores, tests which are not sufficiently reliable, use of tests measuring complex abilities rather than measures of discrete mental traits, a paucity of variables, and the "Purification Technique."

FACTOR LOADINGS DERIVED THROUGH THURSTONE'S SIMPLIFIED MULTIPLE FACTOR METHOD OF ANALYSIS[8]

Thurstone's method of factorial analysis is applicable to a large number of factors and variables and is not dependent upon the absence of group factors for its application. Multiple factor analysis aims to give a more comprehensive analy-

[8] Table of intercorrelations upon which multiple factor analysis and tetrad analysis are based is given in the February 1937 issue of this JOURNAL.

sis than tetrad analysis of the factors involved since the first factor loadings are supposed to give the same information as tetrad analysis.

Table 1 gives the factor loadings for the 18 test variables. Spelling I is not included in factorial or tetrad analysis because of its skewness. Five factors account for most of the relationships between the variables.

TABLE 1

Factor Loadings Determined by the Simplified Method of Thurstone's (Center of Gravity Method)

TEST	I	II	III	IV	V	h^2
1	.468	−.368	.087	.128	.143	.398
2	.310	−.029	−.075	−.506	.118	.373
3	.366	.065	−.214	−.337	.042	.300
4	.432	−.081	−.427	−.053	.195	.417
5	.680	−.030	.281	−.186	−.131	.594
6	.725	−.152	.342	.090	−.182	.707
7	.467	−.218	−.103	.116	−.076	.296
8	.613	−.156	−.320	−.005	.472	.725
9	.465	−.222	−.301	.247	.181	.450
10	.592	−.327	−.274	.145	.119	.567
11	.680	.609	−.196	.237	.063	.931
12	.662	.430	−.083	.167	−.008	.658
13	.578	.039	.292	−.002	.290	.050
14	.612	−.068	.172	−.185	.106	.455
15	.555	−.065	.245	−.067	.116	.391
16	.623	−.123	.368	.274	−.142	.633
17	.606	.262	.116	−.039	−.292	.537
18	.654	.277	.065	.018	−.033	.510
Σk^2	5.8845	1.1255	1.0893	.7262	.5360	9.3615
$\frac{\Sigma k^2}{n}$	.327	.063	.061	.040	.0297	

1. Academic ability
2. Finger dexterity
3. Tweezer dexterity
4. Spatial relations
5. Clerical number checking
6. Clerical name checking
7. Feature discrimination 2A
8. Feature discrimination 3A
9. Feature discrimination 4A
10. Feature discrimination 5A
11. Number cancellation
12. Letter cancellation
13. Letter substitution
14. Number substitution
15. Verifying arithmetic
16. Spelling II.
17. Number group checking
18. Addition

$\frac{\Sigma k^2}{n}$ is the sum of the squares of the factor loadings divided by the number of variables.

h^2 is the communality, which is the sum of the squares of its factor loadings. When h^2 is less than 1.00 it is an indication that the test has either a specific or unique factor or a chance error factor in the score, or both of these.

All of the tests are positively loaded with the first factor. Clerical name checking is most heavily weighted with this factor, while clerical number checking and number cancellation are next most heavily loaded. The clerical tests are most heavily loaded with the first factor, while spatial relations and academic ability receive somewhat smaller weightings. The feature discrimination tests are more heavily weighted with factor one than are spatial relations or academic ability, but less than most of the clerical tests. The two dexterity tests are the most unique in respect to the first factor. Factor one accounts for 33 per cent of the total variance. This factor seems to be measuring a general ability factor, one which emphasizes ability in clerical tests.

The second factor, which accounts for 6 per cent of the total variance, is most highly represented in the number cancellation test with a loading of .61. The letter cancellation test has a weighting of .43 with the second factor. Number group checking and addition have positive second factor loadings. The test which has the highest negative weighting with the second factor is academic ability. This second factor seems to be measuring speed in simple discriminations, an ability which is independent of academic ability for this group.

The third factor also accounts for about 6 per cent of the total variance. After interchanging signs the spatial relations test may be chosen as representing the third factor, with the dexterity, feature discrimination, and number cancellation tests also possessing positive loadings. The two clerical tests, spelling, and arithmetic tests have the highest negative weighting with factor three. The third factor may be best described as a spatial ability factor.

By interchanging the signs in the fourth factor loadings it becomes evident that finger dexterity is most highly saturated with this factor. Other tests which are positively related are tweezer dexterity, and clerical number checking, and those which are negatively weighted are feature discrimination 4A, number cancellation, and spelling. This factor is evidently

speed in motor ability; it accounts for 4 per cent of the total variance.

The fifth factor explains about 3 per cent of the total variance. Here the largest loading is in the feature discrimination test 3A. The clerical checking and the spelling tests are among those which have negative loadings. This fifth factor, though of little importance, seems to represent the ability to observe and compare as measured by the feature discrimination tests, particularly by those of moderate difficulty.

The communalities indicate that number cancellation is adequately accounted for by the five factors postulated. Over 70 per cent of the variance for clerical name checking and feature discrimination 3A are accounted for by these factors; and clerical number checking, feature discrimination 5A, letter cancellation, letter substitution, spelling, number group checking, and addition are fairly well described by the same factors. Tweezer dexterity and 2A feature discrimination are the most unique tests in the battery. The clerical checking tests would not be expected to measure specific factors in this group of tests simply because they are chiefly clerical tests. A later test of uniqueness will be made when the Minnesota Clerical Test will be the only clerical test in the battery.

Factor loadings have been calculated when the clerical tests only were included. Now, if the first factor is a general clerical factor and the second is speed in simple discrimination, then the first two factors should remain rather constant if the analysis is to have much practical value. Variability in the other three factors is to be expected since they are supposedly measuring factors not involved in clerical ability to any great extent.

Table 2 shows that four factors will account fairly well for the total variance. All of the clerical tests have high positive loadings with the first factor which accounts for 45 per cent of the total variance. This factor, as in Table 1, seems to be a general clerical ability factor. Clerical name checking has the highest loading and verifying arithmetic the lowest load-

ing. The change in first factor loadings for the clerical tests from Table 1 to Table 2 are slight, ranging from .01 to .06 with an average discrepancy of .032. This difference is small when it is considered that eight variables out of a total of 18 have been eliminated in the second calculation.

TABLE 2

Factor Loadings for the Clerical Tests Determined by the Simplified Method

TEST	I	II	III	IV	h^2
1	.714	.221	.136	−.420	.753
2	.744	−.292	−.119	.073	.703
3	.713	.498	.183	−.043	.791
4	.722	.432	.478	−.041	.939
5	.613	−.201	.177	−.560	.761
6	.633	−.123	−.061	−.100	.430
7	.579	−.214	.036	.186	.417
8	.659	−.281	.014	.041	.515
9	.663	.193	−.422	.092	.664
10	.681	.215	−.288	.016	.593
Σk^2	4.5424	.8324	.590	.5536	6.5184
$\frac{\Sigma k^2}{n}$	.454	.083	.059	.055	

1. Clerical number checking
2. Clerical name checking
3. Number cancellation
4. Letter cancellation
5. Letter substitution
6. Number substitution
7. Verifying arithmetic
8. Spelling II
9. Number group checking
10. Addition

Number cancellation is the most heavily loaded with the second factor which accounts for 8 per cent of the total variance. Other tests positively loaded with this factor are letter cancellation, clerical number checking, number group checking, and addition. Clerical name checking and spelling gave the largest negative loadings. This factor seems to correspond to the second factor which is present when all of the variables are considered, namely that of speed in simple discriminations. Although the general trend of results is not significantly changed, variation in the test battery does alter the second factor loadings.

Number cancellation is the test most adequately explained by the four factors. Other tests which are well accounted for

by the postulation of these factors are clerical number and name checking and letter substitution. Number substitution and verifying arithmetic are the most unique tests in this battery of clerical tests.

A conclusion to be drawn from this study is that after the withdrawal of a few "outside tests" the main factor will vary little, whereas the factors somewhat less removed will change in the absolute amount of the loadings. The fluctuations in factor loadings depend, then, upon the nature and number of tests removed in the modification. Perry (16) found constancy in the first two factors when psychological and high school content examinations were added to the battery of personality tests.

THE EFFECT OF THE HOMOGENEITY OF THE GROUP ON FACTOR LOADINGS AND SPECIFICITY

In studying uniqueness or specificity of abilities we must consider differences produced by changes in the group of subjects. Tests which may be relatively independent for one group may be more related for other selected groups. If the interrelation of abilities is to be discussed theoretically, it is important for such correlational and factorial analyses to be based upon a homogeneous group. In the practical situation, however, vocational psychologists are not always concerned with very highly selected groups but are often concerned with groups which may be homogeneous only in respect to age or sex. Although it is more difficult, and perhaps dangerous to give any weight to factors which are found under such conditions, it is important to study and compare the factor loadings and specificity of tests for homogeneous and less well selected groups.

The tests selected for such a study include the battery of tests used by the Employment Institute. They are comprised of tests in academic ability, clerical number and name checking, finger and tweezer dexterity, and spatial relations. Since the average score on the Pressey Senior Classification and

Verification Tests was used in this study of clerical workers, the average correlation for these two tests was used for the more heterogeneous groups although the results have been calculated separately for the two tests.

Table III gives the factor loadings on these tests for the

TABLE 3

Factor Loadings for Young Women Clerical Workers on the Institute Battery of Tests

TEST	I	II	III	h^2
1	.473	.170	−.355	.379
2	.693	.286	.361	.692
3	.705	.435	.011	.686
4	.450	−.334	.168	.342
5	.492	−.322	.162	.372
6	.387	−.320	−.258	.319
Σk^2	1.7952	.6176	.3773	2.7901
$\frac{\Sigma k^2}{n}$	.299	.103	.063	

1. Academic ability
2. Clerical number checking
3. Clerical name checking
4. Finger dexterity
5. Tweezer dexterity
6. Spatial relations

original group of 155 young women clerical workers. The first factor postulated accounts for 30 per cent of the total variance. Clerical number and name checking are the most heavily loaded with this factor. Spatial relations has the lowest loading with factor one. This first factor seems to be a general ability factor, one having particular reference to clerical ability.

The second factor, accounting for 10 per cent of the total variance, is most highly represented in clerical name checking. Other tests which have positive loadings are clerical number checking and academic ability. Performance tests, namely, finger and tweezer dexterities and spatial relations have negative loadings with factor two. This factor may be called clerical ability with particular emphasis upon the verbal element in contrast to performance test ability.

Six per cent of the total variance is accounted for by the third factor, which is present in clerical number checking more

than in any other test. Whereas finger and tweezer dexterity have positive loadings with this factor, academic and spatial abilities are definitely unrelated to it. The presence of relationship between dexterity and clerical number checking was also mentioned in the discussion of the intercorrelations.

The three factors postulated account for the two clerical tests rather adequately. Academic ability, spatial relations, and finger and tweezer dexterity tests are measuring rather specific abilities. It is apparent that if only one of the clerical tests was included, that test, too, would be relatively independent.

STANDARD SAMPLE OF UNEMPLOYED WOMEN WORKERS

Darley and Rempel (14) report the intercorrelations for the standard sample of 131 women workers on the Employment Institute battery of tests. In this group the occupational classifications correspond to that for women workers in the population. Darley and Rempel give correlations which show that age is not an important factor in the existing relationships. The average woman worker in the group of unemployed women is thirty years. The average scores are equal to average or C grade abilities (8).

Table IV gives the factor loadings of these tests for the standard sample of 131 unemployed women workers. All of the tests are positively related to the first factor which accounts for 42 per cent of the total variance. Clerical name checking has the highest loading with this factor, clerical number checking and academic ability next. Finger dexterity is the most specific test with respect to loading with the first factor. This factor, as in the previous analyses, seems to be a general ability factor, one emphasizing clerical ability.

The second factor which accounts for 8 per cent of the total variance is most markedly present in clerical name checking. Although the loadings for the first and second factors differ from those for the 155 clerical workers, it is apparent that the same two factors are operating but in varying degrees depend-

ing upon the composition of the group. Academic ability and clerical number and name checking have opposite loadings from spatial relations and finger and tweezer dexterity in the second factor. This factor differentiates clerical from mechanical or motor abilities.

For the standard sample, as for the clerical workers, the third factor which explains 4 per cent of the total variance, is most highly represented in clerical number checking. The

TABLE 4

Factor Loadings for the Standard Sample of 131 Unemployed Women Workers

TEST	I	II	III	h^2
1	.738	.282	−.258	.692
2	.784	.205	.361	.787
3	.840	.349	−.028	.829
4	.315	−.316	.184	.233
5	.466	−.237	−.105	.284
6	.574	−.259	−.041	.399
Σk^2	2.512	.467	.245	3.224
$\frac{\Sigma k^2}{n}$	.419	.078	.041	

1. Academic ability
2. Clerical number checking
3. Clerical name checking
4. Finger dexterity
5. Tweezer dexterity
6. Spatial relations

chief difference is that, for the clerical workers, finger and tweezer dexterity have positive loadings, while for the standard sample only the finger dexterity test has a positive loading. For both groups academic ability is the test which has the highest negative loading. Clerical number checking and academic ability are well accounted for by the three factors postulated. Finger dexterity, tweezer dexterity, and spatial relations are relatively unique.

The intercorrelations for another group were studied by factorial analysis. The data for this group have been taken from the study of Taylor and Darley (14) who report the intercorrelations for a group of 250 unemployed women from various occupational classes. The average age for this group of

workers is twenty-seven years. In this group the average worker has B-grade ability in the various tests, with the exception of finger and tweezer dexterity in which tests she has C-grade ability. This group, except for their superior score in spatial relations, is like the group of women clerical workers in all ability test scores.

Table V, which gives the factor loadings, indicates that three factors account for much of the variance. The first factor, that of general ability, particularly for clerical work, is most markedly present in clerical name checking. All of the variables are positively loaded with the first factor, especially so the clerical and academic ability tests.

TABLE 5

Factor Loadings for a Group of 250 Unemployed Women Workers on the Institute Battery of Tests

TEST	I	II	III	h^2
1	.647	−.280	.223	.547
2	.804	−.205	−.225	.739
3	.842	−.340	−.074	.830
4	.500	.459	−.211	.506
5	.451	.231	−.214	.302
6	.437	.154	.401	.376
Σk^2	2.418	.654	.357	3.428
$\frac{\Sigma k^2}{n}$	.403	.109	.059	

1. Academic ability
2. Clerical number checking
3. Clerical name checking
4. Finger dexterity
5. Tweezer dexterity
6. Spatial relations

Whereas 40 per cent of the total variance may be accounted for by the first factor, about 11 per cent may be attributed to the second factor. The second factor, similar to that one found in the other two groups indicates that clerical number and name checking and academic ability are loaded oppositely to spatial relations, finger dexterity and tweezer dexterity. The difference is that for the clerical workers and the standard sample the positive loadings were in the spatial relations, finger dexterity and tweezer dexterity tests. This factor dif-

ferentiates academic and clerical abilities from performance tests of skill and mechanical ability.

The third factor, which accounts for 6 per cent of the total variance, is most markedly present in the spatial relations test. Only one other test, academic ability, has a positive loading with this third factor. The third factor for this group of workers is a spatial ability factor.

Clerical number and name checking are the most adequately accounted for by the three factors postulated. Academic ability and finger dexterity are fairly adequately described by these three factors. Tweezer dexterity is the most unique test in the battery for this group of subjects.

TETRAD ANALYSIS OF THE INTERCORRELATIONS FOR THE YOUNG WOMEN CLERICAL WORKERS

The application of the tetrad difference criterion will test the plausibility of a single factor interpretation and it should point out the existence of group factors in the data if any exist. Multiple factor analysis, which is unhampered by the presence of group factors, has indicated the necessity of a multiple factor interpretation.

There are several methods for testing the significance of the tetrads in any particular study. If a general factor and specific factors can describe the existing tetrads, the tetrad differences should be equal to or fluctuate about zero, with differences equal to no more than that expected from sampling errors. The tetrad differences may be considered equal to zero and fulfilling the criterion when:

1. the median tetrad difference and the probable error of the tetrad differences are about equal
2. the largest tetrad is less than four (some writers say five) times the magnitude of its probable error
3. the actual probable error of the observed tetrad differences is equal to or less than the theoretical probable error of a purely chance distribution of such tetrad differences.

The first group of tests analyzed were those having the highest first factor loadings when all of the tests were included in the analysis. There are nine tests in this group, including clerical number and name checking, feature discrimination 3A, number and letter cancellation, number substitution, spelling II, number group checking, and addition. These nine tests with factor loadings between .60 and .73 necessitate the calculation of 378 tetrads.

These tetrads do not fulfill any of the requirements for the assumption of a single general factor. The median tetrad is .044 while the probable error of the distribution of tetrads is .064. These data indicate the presence of group factors in that the probable error of the observed tetrads of .064 is larger than the theoretical probable error of .0273. The probable error of the largest tetrad, .4154, is .0482, calculated by the Spearman-Holzinger Formula and .0463 by the Kelley Formula. Many of the tetrads are more than five times the size of this probable error.

These results indicate that the data are made up of tests which will not satisfy the tetrad criterion. When the various bonds are studied, the evidence points to the existence of four group factors, one between clerical number and name checking, a second between clerical name checking and spelling II, a third between number and letter cancellation, and a fourth between number group checking and addition. Group factors may be controlled by omitting them from the battery or by combining them into one coefficient. Various techniques were used to verify the postulation of group factors.

The first attempt to eliminate certain group factors was made by dropping clerical name checking, number cancellation, and number group checking from the battery of tests. The six tests remaining were clerical number checking, feature discrimination 3A, letter cancellation, number substitution, spelling, and addition. These tetrads seem to fulfill the tetrad criterion for the assumption of a general factor. The median

tetrad is .019 while the probable error of the observed tetrads is .0206. Another indication of the elimination of group factors is that the probable error of .0206 is less than the theoretical probable error of .0267. The tetrads are then within the boundaries allowed for sampling errors. None of the tetrads are significantly greater than their probable errors. By the elimination of three of the tests which were closely linked to other tests, group factors were successfully eliminated.

Since these tests seem to fulfill the criterion, the correlation between each test and the 'G' factor has been calculated, with the following results:

Clerical number checking	r1g	.701
Feature 3A	r2g	.570
Letter cancellation	r3g	.692
Number substitution	r4g	.628
Spelling II	r5g	.594
Addition	r6g	.665

These results correspond to the findings of multiple factor analysis which also showed that clerical number checking and letter cancellation had high loadings with factor one, while feature discrimination 3A and the spelling test were loaded less heavily.

Another combination was used in studying the tests possessing high first factor loadings. An attempt was made to eliminate group factors by dropping number cancellation, clerical number checking, letter substitution, spelling, and number group checking. The tests remaining are clerical name checking, feature discrimination 3A, letter cancellation, number substitution, and addition. Group factors have not as successfully been eliminated from this group of tests. The median tetrad is .0325 which is greater than the probable error of the tetrads, which is .0284. The observed probable error of the tetrads is .0284, which is greater than the theoretical probable error of .0265 that is allowed for sampling errors. None of the tetrads, however, are significantly larger than their probable

errors. When the tests are correlated with the general factor, feature discrimination again is the least related. The correlations are:

Clerical name checking	r1g	.647
Feature 3A	r2g	.573
Letter cancellation	r3g	.683
Number substitution	r4g	.643
Addition	r5g	.663

Spearman's technique of eliminating group factors by the combination of two series was used in this study. The tests were grouped as follows: clerical number and name checking, feature discrimination 3A, number and letter cancellation, number substitution, and number group checking and addition. These tetrads fall within the range allowed for sampling errors. The median tetrad is .0181 which is about equal to the probable error of .0153 for the tetrad distribution. These tetrads fulfill the criterion for a single general factor and specific factors, in that the observed probable error is .0153 which is less than the theoretical probable error of .0274 allowed for sampling errors. This technique as well as the method of elimination of the variables substantiates the postulations of group factors in the original group of tests studied. The correlation of each test with 'G' is as follows:

Clerical number and name checking	r1g	.764
Feature 3A	r2g	.555
Number and letter cancellation	r3g	.726
Number substitution	r4g	.665
Number group checking and addition	r5g	.729

Various other combinations have been studied by means of factorial analysis. One of the groupings is clerical number and name checking, number and letter cancellation, number substitution, and number group checking and addition. The tetrads for these tests are nearly zero, indicating that they fulfill the criterion.

Another test combination which was studied consisted of clerical number and name checking, feature discrimination 3A, number and letter cancellation, and number group checking and addition. The tetrad differences for these test combinations also fulfill the criterion for a single general factor.

When factor analysis techniques are applied to tests of number and letter cancellation, number group checking, and addition, tests which were heavily loaded with the second factor, the results indicate the presence of a group factor in the cancellation tests, one in number group checking and addition, or one in each of these pairs of tests.

Spatial relations and feature discrimination 3A, 4A, and 5A are the tests which have the highest third factor loadings. The tetrad differences for this group are approximately equal to zero, which indicates that a central factor plus specific factors may account for these relationships.

The fourth factor is most highly represented in finger and tweezer dexterity, clerical number checking, and number substitution. The tetrads are all insignificant if the criterion of five instead of four times the probable error is accepted as the boundary of statistical significance.

The heaviest fifth factor loadings are in feature discrimination 3A and 4A, letter substitution, and spatial relations. The tetrads for these tests are equal to zero and satisfy the criterion.

A tetrad analysis has been made of the relationships between the verbal-clerical tests. The tetrads are very close to zero, indicating that the clerical name checking, letter cancellation, letter substitution, and spelling II tests fulfill the requirements for a single general factor. Correlations between 'G' and the test variables are for:

Clerical name checking	r1g	.770
Letter cancellation	r2g	.486
Letter substitution	r3g	.579
Spelling II	r4g	.816

These correlations indicate that the factor measured by these tests, which may be called a verbal-clerical factor, is most adequately represented in the spelling test. The second highest relationship is between name checking and 'G.' These results are what might be expected from the correlations between academic ability and these clerical tests. Spelling has the highest relationship with academic ability and cancellation the lowest.

Six tests have been classified under the title of numerical-clerical tests. These tests include clerical number checking, number cancellation, number substitution, verifying arithmetic, number group checking, and addition. The average intercorrelation for this group of tests is .422. The tetrads calculated for these intercorrelations do not satisfy the tetrad criterion. The median tetrad is .039 while the probable error of the tetrad distribution is .0447. Another indication of the lack of conformity to the criterion is the fact that the probable error of the actual distribution of tetrads is .0447 which is larger than the theoretical probable error of .0272. These data indicate the presence of group factors. A study of the possible bonds indicates the presence of group factors in clerical number checking and arithmetic, and in number group checking and addition.

To test this postulation of group factors, these two groups, addition and number group checking, and clerical number checking and arithmetic were combined into two measures. The tests are now, number cancellation, number substitution, clerical number checking and verifying arithmetic, and number group checking and addition. Tetrads for these test combinations are within their probable errors and fulfill the criterion. Correlations between the central factor and these four tests indicate that number group checking and addition are determining 'G' more than any other test combination. The correlations are for:

Number cancellation	r1g	.617
Number substitution	r2g	.416
Clerical numbers and arithmetic	r3g	.659
Number group checking and addition	r4g	.780

A tetrad analysis has been made of clerical name checking and "outside tests," including academic ability, finger and tweezer dexterity, spatial relations, and feature discrimination 4A. The median tetrad for this distribution is .045 and the probable error, .0380, whereas the theoretical probable error is only .0210. The largest tetrad is on the borderline of significance, in that its probable error is .0281. These data seem to indicate the possibility of a minor group factor, which may exist in clerical name checking and academic ability.

Clerical number checking and the "outside tests" have also been analyzed by tetrad analysis. The tests include clerical number checking, academic ability, finger dexterity, tweezer dexterity, spatial relations, and feature discrimination 4A. The median tetrad for this distribution is .047 and the probable error .0335. Another tetrad criterion is not satisfied in that the observed probable error is .0335 which is larger than the theoretical probable error of .0210. The largest tetrad is only .0890, yet the evidence does not point with certainty to the existence of a general factor for these tests.

The four feature discrimination tests have been studied by factor analysis techniques. These tetrads border on significance in that the probable error for the largest tetrad is .0384 or .0379, depending on the formula. If a tetrad four times the probable error is accepted as significant, the tetrad .1625 is significant, but it is not if the criterion of five times the probable error is accepted. The group factor, if present, undoubtedly lies in the feature discrimination tests 4A and 5A. Assuming, however, that they satisfy the criterion, we find the following correlations between 'G' and the variables:

Feature discrimination 2A	r1g	.526
Feature discrimination 3A	r2g	.653
Feature discrimination 4A	r3g	.652
Feature discrimination 5A	r4g	.631

These correlations agree with results reached by multiple factor analysis, in that both indicate that feature discrimination 2A is the most unique test in the battery of feature discrimination tests.

Tetrad differences have been computed for clerical name checking, spelling, clerical number checking, and verifying arithmetic, which tests seem to represent best the verbal and numerical elements. These tetrads indicate that more than one factor must be postulated to account for these intercorrelations.

Tetrad analysis of clerical name checking, spatial relations, spelling, and feature discrimination 5A also results in tetrads which do not satisfy the criterion for a single general factor.

When tetrad differences are calculated for clerical number checking, spatial relations, verifying arithmetic, and feature discrimination 5A, the results again indicate failure to meet the tetrad criterion for the postulation of a general factor.

The above three comparisons show, when the variables contain two pairs of tests, each pair being closely related and measuring a different ability from the other pair of variables, that the tetrads indicate group factors and the independence of the factors involved in each group. Anastasi (1) finds similar results when she compares verbal, memory, or numerical tests in the same group of variables.

On the basis of tetrad analysis, it is possible to select out five tests from the original Institute Battery which seem to be relatively unique. These five tests which do not overlap are Pressey Education, clerical number checking, finger dexterity, tweezer dexterity, and spatial relations. The median tetrad for this distribution is .045 which is somewhat larger than the probable error of .0279. Although none of the tetrads are large, the theoretical probable error is only .0206 while the

observed probable error is .0279. Neither of these two criteria is satisfied. The largest tetrad in the distribution is not significantly greater than its probable error of .0207. If the assumption is made that the criterion has been satisfied, the correlations between the general factor and the test variables are for:

Pressey education test	r1g	.332
Clerical number checking	r2g	.445
Finger dexterity	r3g	.482
Tweezer dexterity	r4g	.574
Spatial relations	r5g	.429

For this homogeneous group of clerical workers, the tweezer dexterity test has the highest relationship with general ability. Multiple factor analysis has shown, when applied to a similar battery of tests with the addition of clerical name checking, that at least three factors can be postulated to account for these relationships. The factors which the multiple factorists call their second and third factors are called group factors by the single factorists.

Flanagan (7) stated in his critical review of factor analysis, that "it would be a contribution to the two factor position if tests, which by a posteriori selection, are decided to be definitely in accordance with the criterion of zero tetrads were applied to new groups and shown to fit the criterion with allowance for sampling error." Assuming that the above tests approximately satisfied the tetrad criterion, a check will be made to see if these same tests when administered to a different group approximate the criterion. Since it was not possible to match the clerical group in homogeneity, a group of 131 women in the standard sample, (14) which was drawn from various occupations, was selected. Even though the results cannot be used without caution, because of this heterogeneity, they will point out the changes in the tetrads and in the weights of the variables in determining the central factor when the nature of the group is radically altered. The median tetrad is .038 and the probable error is .034. The observed probable error is some-

what larger than the theoretical probable error of .0244. These tetrads seem to satisfy the criterion in that the largest tetrad is not significantly greater than its probable error, .0302. Assuming that these tetrads can be accounted for by a single general factor and specific factors we find the correlations between each variable and the central factor are:

Pressey education test	r1g	.588
Finger dexterity	r2g	.316
Tweezer dexterity	r3g	.467
Spatial relations	r4g	.631
Clerical number checking	r5g	.690

These correlations differ from those for clerical workers and indicate that, similar to multiple factor analysis techniques, the results obtained depend to a considerable extent upon the nature of the group. Whereas the tweezer dexterity test measures the central factor more adequately than do the other tests in the clerical group, the clerical number checking test has the most weight in determining the general factor for the standard sample. Academic and spatial ability have larger weights in determining the central factor for the standard sample than for the clerical workers, while finger and tweezer dexterity have heavier weights for the clerical workers than for women in the standard sample.

The study of these two groups by Thurstone's Technique of Multiple Factor Analysis showed exactly the same trends, finger and tweezer dexterity being more unique for the standard sample than for the clerical workers, while the academic, clerical, and spatial relations showed greater specificity for clerical workers than for the standard sample of women which is a much more heterogeneous group.

Tetrad analysis of these five tests for the two groups has shown that the tests are measuring five abilities which are related but do not overlap enough to indicate the presence of group factors.

CONCLUSIONS

1. Thurstone's Multiple Factor Method is a valid technique in that the nature of the main factors is relatively constant even with modifications in the test material and in the make-up of the group. But the precise amount of loadings for each test is variable and dependent upon changes in the test battery and in the homogeneity of the group. The number and amount of changes in factor loadings increases as the groups or tests become increasingly dissimilar from the original set-up.

2. The results indicate that all of the clerical tests are positively loaded with one factor and that although the Minnesota Clerical Test is positively related to other clerical tests it is measuring a specific ability which is relatively independent of spatial, academic, and dexterity abilities. As the group becomes more heterogeneous, however, the clerical test becomes more closely related to academic ability. This change may be detected in the first factor loadings for these three tests as well as in their communalities. As the groups become less homogeneous, the clerical and academic ability tests tend to become less specific.

3. When the results from factorial analysis are considered together with results gathered from the application of the clerical tests to the differentiation of high from low grade clerical workers of good from poor commercial students, and of clerical workers from workers in other occupations, the clerical test seems to be sufficiently valid and specific to justify inclusion in a vocational testing program. Although correlational and factorial analysis indicate that clerical number and name checking tests are in part measuring the same abilities, yet other results have indicated that the inclusion of both of the tests yields the maximum validity. Factorial analyses indicate that the other tests of the Employment Institute Testing program are sufficiently unique, reliable, and valid to be included in such a testing program.

Multiple factor analysis is a valuable statistical technique in that it is a means of discovering the specificity of the tests being

used. In all practical applications, its dependence upon changes in the test battery and upon the nature of the group must not be overlooked.

4. Tetrad differences for the tests with the highest first factor loadings, when the eighteen variables are considered, did not fulfill the criterion but indicated the presence of group factors between clerical number and name checking, clerical name checking and spelling, number and letter cancellation, and between number group checking and addition. With elimination of the overlapping by combination or removal of the disturbing tests the tetrad criterion was satisfied. A study of all of the tests dealing with the numerical element indicates the presence of another group factor between clerical number checking and arithmetic. The verbal clerical tests and the numerical groups fulfill the tetrad criterion. The feature discrimination tests were found to include a group factor.

5. Some of the test combinations satisfy the criterion while others indicate the presence of group factors or fail to meet the criterion. A tetrad analysis was made of the Institute battery of tests for the homogeneous group of clerical workers and for the standard sample of women workers. The results indicate that the individual tetrads are not significant and that academic, spatial, finger, and tweezer dexterities, and clerical number checking are measuring abilities, which though related, are relatively unique.

When changes in tetrads are noted for the homogeneous clerical and the heterogeneous standard sample groups, the same results are found as when factor loadings were compared for the two groups. With increase in heterogeneity, academic, clerical, and spatial abilities become less specific while finger and tweezer dexterity become somewhat more specific. As for multiple factor analysis, change in the group tested altered the test weights in determining the central factor.

6. When multiple factor and tetrad analyses are applied to the same data, the results are similar but multiple factor analysis gives results which are more complete with the expenditure

of less effort. Tetrads, as Thurstone has stated, involve laborious procedure which gives no more than what is obtained by the first factor loadings. The fact that all correlations are positive indicates the existence of a general factor, while relatively high correlations indicate the presence of group factors, or in other words, overlapping.

7. This study has shown then that the Minnesota Vocational Test for Clerical Workers is measuring an aptitude which is related positively to the abilities to observe and compare, to discriminate small differences rapidly, to adjust to a new situation and to give attention to a problem. Although the clerical checking tests are closely related, clerical number checking measures more of a numerical factor and clerical name checking more of a verbal factor. The clerical test has shown itself to be relatively unique with respect to academic, spatial, and dexterity abilities. As the groups become more heterogeneous, however, academic, clerical, and the spatial tests become less unique.

8. Factor analysis, particularly multiple factor analysis, has proved a valuable method in the analysis of abilities in a practical situation. Such a method reveals something of the nature of the specificity and the variations in the specificities of the tests with changes in the group studied. That even the general trend of results is relatively constant with changes in the variables or group studied is evidence for the validity of factor analysis techniques. Changes which alter the correlation coefficients upon which the factors are based would be expected to bring about corresponding variations in the size of the factors.

9. One of the major aims underlying this project was to discover whether or not tests tapping a few broad abilities could be substituted for many tests of specific abilities in the clerical field. It was known at the outset that the tests were reliable and valid measures of clerical ability, but it was not known what other tests, not specifically clerical, were measuring abilities similar to those involved in the Minnesota Clerical Test.

Ideally, the vocational guidance counselor desires the fewest number of tests to measure the maximum number of independent abilities important for success in a given vocation.

10. The question, therefore, was raised regarding the reduction of the number of tests used as a battery in a definite comprehensive vocational guidance program, that of the Employment Stabilization Research Institute. As pointed out frequently, this battery of tests included nine tests designed to sample many forms of skill of possible significance in clerical, sales, and industrial occupations. Only six of the original nine tests were used in this analysis.

11. Analysis indicates that only one of the tests studied, that of clerical name checking, if necessary, could have been eliminated without seriously interfering with the value of the battery. This test overlaps clerical number checking and for heterogeneous non-clerical groups it also overlaps academic ability. Clerical ability, then, like mechanical ability (15), involves overlapping group factors. Group factors are present between clerical number and name checking, clerical name checking and spelling, clerical number checking and arithmetic, clerical name checking and academic ability, the two cancellation tests, and between number group checking and adding. Multiple factor analysis indicates the necessity for the postulation of a minimum of four factors to account for the interrelations among the clerical tests. The inclusion of both clerical number and name checking can be defended. Thurstone states that if a battery of tests is valid, it is a "good one, quite irrespective of whether it requires more than one factor to describe the intercorrelations. There is nothing intrinsically good or bad about the presence of more than one common factor" (23). Further reason for the retention of both of the clerical checking tests in the battery is that the validity of the test is considerably increased when both of the tests are included (3). Multiple factor analysis indicates that a common factor runs through all of the tests and in addition there are minor group factors.

12. If a more limited battery is desired for clerical workers alone, then the analysis undertaken in this paper proves the inutility of finger dexterity, tweezer dexterity, and the spatial relations tests, and the great value of tests of academic and clerical abilities. This selected group of clerical workers was not differentiated from women workers in the standard sample by the dexterity and spatial tests. Although statistical analysis shows these tests to be measuring very specific traits, they are not important in the prognosis of clerical ability.

Should circumstances demand economy of time and effort and yet the most adequate prediction and guidance possible, the academic ability test and the clerical number checking test will serve the purpose. If conditions permit the use of only one test, then the most valuable one for the prediction of clerical success is clerical name checking, in that this test overlaps both academic ability and clerical number checking.

These conclusions issue out of the inspection of data available as a result of the process of factor analysis and the less technical measures of validity and reliability.

REFERENCES

1. ANASTASI, A. "A Group Factor in Immediate Memory," *Arch. Psychol.*, 1930, No. 120.
2. ————. "Further Studies of the Memory Factor," *Arch. Psychol.*, 1932, No. 142.
3. ANDREW, D. M., and PATERSON, D. G. Measured Characteristics of Clerical Workers, Bull. of Empl. Stab. Res. Inst., Univ. of Minnesota, 1934, Vol. III, No. 1.
4. BROWN, W., and STEPHENSON, W. "A Test of the Theory of Two Factors," *Brit. J. Psychol.*, 1933, 23, 352–370.
5. BURT, C. Distribution and Relation of Educational Abilities, 1916, London County Council, p. 46.
6. CARTER, H. D. "A Reply to Prof. Spearman," *J. Educ. Psychol.*, 1930, 21, 627–628.
7. FLANAGAN, J. C. Factor Analysis in the Study of Personality, Stanford Univ. Press, 1935.
8. GREEN, H., BERMANN, I. R., PATERSON, D. G., and TRABUE, M. R. A Manual of Selected Occupational Tests for Use in Public Employment Offices, Bull. of Emp. Stab. Res. Inst. Univ. of Minnesota, 1933, Vol. II, No. 3.

9. GREENE, E. B. Manual for the Michigan Non-Verbal Series, Ann Arbor, Edwards Bros., 1931.
10. HOLZINGER, K. J. Statistical Resume of the Spearman Two Factor Theory, Chicago, Univ. of Chicago Press, 1930.
11. JORGENSON, C. "Analysis of Some Psychological Tests by Spearman Factor Method," *Brit. J. Educ. Psychol.*, 1934, 4, 96–108.
12. KELLEY, T. L. Crossroads in the Mind of Man, A Study of Differential Mental Abilities, Standard Univ. Press, Stanford Univ., 1928.
13. LINE, W., ROGERS, K. H., and KAPLAN, E. "Factor Analysis Techniques Applied to Public School Problems," *J. Educ. Psychol.*, 1934, 25, 58–65.
14. PATERSON, D. G. (Editor) Research Studies in Individual Diagnosis, Univ. of Minnesota, Bull. of Emp. Stab. Res. Inst., 1934, Vol. III, No. 4.
15. PATERSON, D. G., ELLIOTT, R. M., ANDERSON, L. D., TOOPS, H. A., and HEIDBREDER, E. Minnesota Mechanical Ability Tests, Minneapolis: Univ. of Minnesota Press, 1930.
16. PERRY, R. C. A Group Factor Analysis of the Adjustment Questionnaire, Univ. of Southern California Educ. Monog., 1933, 1934, No. 5.
17. SCHILLER, B. "Verbal, Numerical, and Spatial Abilities of Young Children," *Arch. Phychol.*, 1934, No. 161.
18. SCHNECK, M. M. R. "The Measurement of Verbal and Numerical Abilities," *Arch. Psychol.*, 1929, No. 107.
19. SPEARMAN, C. The Abilities of Man, London: Macmillan, 1927, 339–366.
20. STEPHENSON, W. "Tetrad Differences for Verbal Sub-Tests Relative to Non-Verbal Sub-Tests," *J. Educ. Psychol.*, 1931, 22, 334–350.
21. ———. "Tetrad Differences for Verbal Sub-Tests," *J. Educ. Psychol.*, 1931, 22, 255–267.
22. ———. "Tetrad Differences for Non-Verbal Sub-Tests," *J. Educ. Psychol.*, 1931, 22, 167–185.
23. THURSTONE, L. L. A Simplified Multiple Factor Method and an Outline of the Computations, Chicago: Univ. of Chicago Press, 1933.
24. ———. "The Vectors of the Mind," *Psychol. Rev.*, 1934, 41, 1–32.
25. WILSON, J. H. "The Nature of Intelligence," *J. Educ. Psychol.*, 1931, 22, 20–34.

FACT AND FANCY REGARDING DRIVER TESTING PROCEDURES

A. R. LAUER
Iowa State College

INTRODUCTION

IN justice to those who desire to use the newer methods of measuring characteristics of automobile drivers, as well as to the tests themselves, it should be stated that the work is still in the experimental stage. The National Research Council has sponsored researches on the factors necessary for safe driving since 1928 but has not as yet recommended specific measures to be used in drivers' examinations. More has been found out regarding the limitations of tests conventionally used and ways of improving the driver's license examinations than of the specific nature of tests to be used for identifying accident-prone drivers.

Driving tests are not new in spite of the claims of would-be inventors who have literally copied earlier pieces of apparatus and passed them off as original. Münsterberg early developed certain tests of functions related to street car manipulation. Viteles and O'Rourke used various forms of driver-testing equipment several years ago. J. M. Lahy, in Paris, has utilized similar tests in selecting street car and bus operators for at least 15 years. Marbe and Lippman have used mechanical tests for classifying drivers in Germany. Kaffka has also conducted experimental work on the judgment of speed and distance. Mira, in Spain, has interested himself in studies of automotive driving. The National Institute of Industrial Psychology in England under C. S. Meyers and others has sponsored such tests and studied many of these problems experimentally.

In 1928 the writer, as a graduate student under Dr. A. P. Weiss, at the Ohio State University and through grants from the National Research Council, made a survey of all available studies up to that time and proceeded to develop ways and means of measuring characteristics of automobile driving. Inquiries were made of those who had done work in the United States, Canada, Great Britain, France, Germany, Italy, Denmark, Holland, Sweden, Japan, China, Australia and South America. Replies were received from about half of these places describing the work which had been done.

STATUS OF TESTING PROCEDURES IN 1928

A digest of this material showed several distinctive characteristics of most earlier studies:

1. There was considerable indication of sporadic attempts made to test drivers of various types, usually for the purpose of meeting some immediate practical situation. The investigators became greatly interested for a time but eventually gave up the work for other problems.
2. Very little basic research had been done on driving ability as such. No definite ground work had been established from which to proceed. At that time a series of studies was outlined and begun at the Ohio State University for the purpose of investigating fundamental factors of automotive manipulation.

ASSUMPTIONS MADE IN EARLY STUDIES

As a basis for setting up experimental studies, certain assumptions had to be made. These assumptions were:

1. That the most fruitful approach to problems of highway safety would be through experimental studies of the traits, response patterns and skills related to driving.
2. That successful driving performance is dependent upon a minimum level of ability or abilities. Below this level accident risk becomes greater and more accidents will inevitably result.
3. That a consistent measurement of such characteristics will give some indication of one's ability to maneuver an automobile.

4. That such traits or characteristics of driving are only a part of the general pattern of responses necessary for safe driving. Many other factors relating to automobile and highway design, traffic regulations and practice, as well as to training in driving, are of equal importance for public safety.
5. That certain existing measures of human traits are well standardized and require only application to the problem at hand.
6. That the fundamental characteristics of psychological testing must be observed to obtain satisfactory results. These are:
   a. The test must simulate as nearly as possible the nature of the responses or trait to be measured.
   b. The measures made must be reliable or consistent in the results obtained.
   c. The measures made must be valid when checked against a genuine criterion.
   d. That measures used for scientific study should be as simple as possible. The addition of "showy" exteriors will not make a poor test valuable.

With this background and acting upon these assumptions a series of cooperative studies was begun and published three years later as a monograph on, "Psychological Principles in Automotive Driving." A part of these studies were carried on in Columbus, Ohio. Cooperating agencies in Dayton, Ohio, and East Lansing, Michigan, were very helpful in forwarding certain phases of the work which could not be developed at the central experiment station.

In 1931 a part of this work was transferred to Iowa State College where others interested in the problem have aided in various studies which have resulted in about twenty-two publications.

As early as 1930, Purdue University had conducted tests of riding comfort in relation to fatigue. These studies have since been articulated with the activities of the Committee on "Psychology of the Highway" of the National Research Council.

Various related studies on the effects of alcohol, psychiatric conditions of accident-prone individuals and fatigue have been

in progress at Milwaukee, Pittsburgh, Cleveland, Detroit, and Chicago respectively.

In 1933 an FERA project was begun at Massachusetts State College and has progressed along the line of improving instruments since that time until it was taken over by the Harvard Bureau of Traffic Research in 1935–36. Some of the original equipment used at Columbus was borrowed and revised by those carrying on the work. The basic principles of these tests are old and will be found described in the literature of the past 25 years. With considerable expense and by using good artists and craftsmen, some spectacular pieces of apparatus have been built. For practical uses, equipment of this type can be constructed for a small fraction of the cost of such elaborate apparatus. In one case a room full of apparatus, costing $300 to $400, has been reduced to an ordinary cardboard 15 inches by 15 inches costing 50 cents.

PRESENT STATUS OF THE STUDIES

A recent writer on this subject in a popular magazine inferred that results obtained from this problem were only one-sixth better than chance in the prediction of accidents. In fairness to those who have contributed to these studies, a few explanatory statements should be made along with the presentation of known facts and limitations of the studies. In the first place, the statement referred to was based upon a relatively small sampling of data collected nearly four years ago. The word *predict* should be supplemented by the word estimate. No large commercial enterprise will *predict* that its income will be a certain figure for the next two years but it can *estimate* the receipts for the coming year at a figure which will mean a substantial profit, subject to a margin of error in either direction. So long as the actual variation of income does not exceed the margin of error allowed, the concern will continue to profit and progress will be made. To resort to *predictions* that specific results may be obtained—as 100,000 shirts will be sold on the 10th day of April of next year—is not only unnecessary but quite out of the realm of the true science of economics.

In other words, improvement in any field will stand on its own merits. To reject that which is not perfect would be to reject most all civilization and scientific progress. Organized society is run on the premise that improvement, even in a minor way, is desirable. Practically all the affairs of life are run on this basis. Otherwise all scientific and technical effort is fruitless. We buy life insurance knowing that the XYZ insurance company lost $50 an acre on a certain farm mortgage in a middle western state last year. The law of averages determines the ultimate pragmatic value of an innovation in the same way that it determines the safety of an investment.

After almost nine years continuous study of the psychological problems of driving there are certain definite facts which may be given with as much certainty as the fact that cod-liver oil is good for rickets. No one would venture to say that the absence of cod-liver oil or its counterpart will prognosticate rickets or that all cases of rickets are amenable to this treatment. It can only be said that the disease seems to be relieved in many instances by this remedy.

The studies carried on by the National Research Council on the psychological factors of automobile driving have been of various types and have by no means been concentrated upon construction of "tests" as such. "Tests" have been one of the by-products of the major investigations but have come in for the greatest amount of public attention. With this thought clearly in mind let us proceed to discuss some of the facts which have been noted during the past nine years of fundamental research. These facts have experimental data to support them and the conclusions are presented impartially.

Newer developments will undoubtedly change these interpretations as is true of all scientific work. For example, medicine reserves the right at any time to change its methods of treatment when a more efficacious remedy is found. Results are never guaranteed. One of the earmarks of charlatanism is a ready-made formula or panacea for all ailments.

SOME FACTS RELATING TO ACCIDENT FREQUENCY

Over a ten year period of driving, various studies have shown that between 20 and 30 per cent of ordinary or lay drivers have all of the accidents in a given population. With commercial companies over a fifteen year period, a study of 1014 drivers showed that 43 per cent had all the accidents. Considering this period of time and numerous non-linear relationships which have been found, it seems that the accident group may be divided into two categories: (a) a small percentage—perhaps 2 to 4 per cent—who are accident-prone, and (b) the remainder of the 20 to 30 per cent who may be designated as accident-liable. The latter term is used with reservation, meaning those who need not have accidents but who for one reason or another do have a few mishaps. Usually these accidents are experienced early in life after which there is a relatively long accident-free period of driving.

Color blindness has long been a suspected cause of accidents. Railroads bar color-blind persons from regular employment. The army and navy have strict requirements for color vision. Our studies have shown that only about one out of six of the worst color-blind persons in any population may confuse colored lights. This means that something less than one-third of one per cent of the population will ever miss a stop light through faulty color perception. Actually, a check against accident records shows no relationship at all between color blindness and accidents. Not a single accident which can be traced to color blindness has come to the attention of the writer during the past nine years.

Activity of the individual as indicated by accepted measures of this characteristic, as well as strength, do show a positive relationship to safety. The relationship is not high, due to certain weaknesses in the criterion as discussed below.

Simple reaction time as such, irrespective of the columns written about it, does not seem to be correlated with accident frequency. In spite of validity coefficients of nearly zero, several exploiters of "driving tests" are basing practically their

whole experimental approach upon one trait. It is not inconceivable that complex reaction time may show different results but this has not been scientifically demonstrated and there is reason to believe the results would not be highly significant. The variability in successive reactions is of more diagnostic value than absolute reaction time *per se*.

There is evidence to support the statement that marked ocular dominance is a factor in safe driving. No clearly defined reason for this can be given at present. Visual acuity below 40 per cent seems to be associated with a higher accident rate and the curve of relationship drops off somewhat for acuity below 60 per cent.

"Safe driving" and what many people call "good driving" are not the same. Many safe drivers are not good drivers; that is, skillful in the sense of mechanical and general alertness. A zero correlation was obtained between company rating on driving ability and accident records of their drivers. The racing driver is usually a "good driver" but we would hardly call him a "safe driver" on the race track. He may be a safe driver when on the road, since many of these men drive at a conservative rate of speed when not making time records.

For reasons enumerated, it is very difficult to separate superior performance records from safe driving records. The techniques so far developed in our laboratory stress four aspects of performance of a complex nature. These are:

(a) Manipulation of controls as such—a form of eye, foot and hand coordination.
(b) Observation. The correct interpretation of and reaction to serially presented stimuli. The emphasis here is upon correct interpretation rather than speed of interpretation within limits of one-half second duration. This is longer than the average reaction time.
(c) The time taken for the performance of a given task. All persons do not work efficiently at the same rate of speed. Each is allowed to set his own rate. It may be considered as a measure of caution.
(d) Observation time. When new stimuli are presented, a certain period of time elapses before they are noticed.

This is much more important than reaction time as such. It is measured in one-fourth second units by the apparatus used in our investigations.

Experimentation is continuing on these factors but it seems clear at present that the latter two are equally as important in evaluating driving performance, even though the highest first order relationships have been found between accidents and observation. This result has been obtained both in outdoor field tests and upon laboratory tests simulating actual performance.

Considerable data have been obtained on hearing but this seems to be important only at extreme ranges. Any relationship which has been found could easily be traced to some associated phenomenon.

High and low blood pressure seem to be somehow related to accident frequency. An explanation of this relationship has not been found for high blood pressure. Systolic blood pressures above 175 should be carefully scrutinized. Those with low blood pressure seem to go to sleep easily and lack the vitality necessary for alertness in long continued driving. The lower limits are around 96–100. Age must be considered.

Several other factors which seem important at the present time and which are being subjected to experimental validations are: (1) perseveration, (2) attitudes toward law enforcement and the public, (3) knowledge and understanding of traffic hazards and safety principles, (4) compensation, and specific attitudes commonly grouped under the heading "recklessness."

It seems quite clear that any validation of test results which does not take into consideration compensation is likely to turn out negatively, although there may be numerous other explanations of negative results.

In this connection, the various results obtained up to the present suggest the fact that accidents are possibly preceded by existing psychophysical conditions which might be called major and minor susceptibilities. One major susceptibility may have as much weight as several minor susceptibilities. Again, one

minor susceptibility may not be operative except in the presence of another minor susceptibility. For example, clumsiness at the wheel may be a minor susceptibility. Coupled with nervousness, it may become a serious handicap while alone it may be easily compensated for by careful driving.

This leads to a further conclusior that the ultimate method of evaluating driving ability will probably be that of identifying patterns of response which may prognosticate accident susceptibility. Several different types of ability-capacity constellations may be identified which will probably be most useful in diagnosing or prognosticating accident-proneness. The concept is not new in that it is similar to that of syndromes identified with certain diseases. Until this is done no highly valid method of picking out accident-prone individuals from a random sampling of the population can be made by any simple test. The use of such methods, when properly administered, may have a certain effect on accident reduction through focusing attention on weaknesses in the same way that a physical examination will tend to safeguard the health of the individual by warning him of ways in which he may protect himself.

By interpreting the results of test data in this manner, a great amount of good may be accomplished. If too much weight is given testing, it may not prove as satisfactory as those using it have anticipated. Specially trained assistants are necessary and the devices must be used with considerable care. Allowance for various psychological differences must be made. The methods must be considered an aid rather than a final court of appeal in drivers' license examining. They are of no more value in the hands of a layman than a dentist's tools are of value to one not trained in dentistry. The important thing to be considered is the training of the person in charge of the equipment. It is a special branch of psychology as remote from conventional routine in graduate training as the techniques used in studying the conditioned reflex are different from those used in the perfection of a personality inventory. This fact must be frankly realized if psychology is to be an asset in solving the accident problem.

One of the difficulties facing those interested in validation studies is that of establishing a criterion. Causes of accidents are no more alike than the causes of deaths. Various and sundry antecedents are to be identified before associated response patterns may be determined. High mileage drivers are found to have a lower accident rate when units of time are equal. Experience, hazards, purpose of driving, time of day, conditions of weather and the roadway, type of vehicle, attitude of the driver, the degree of blame, the seriousness of the accident, manner in which the accident took place, and numerous other variables make the establishment of a criterion extremely difficult.

In spite of the limitations of the methods used and the results obtained, the only scientific attitude one may assume is that research should be continued until results are forthcoming. We do not abandon our medical schools because persons still become ill and die and we do not give up aviation because it is not one hundred per cent efficient or safe. We do not cease the quest for better methods of education for the reason that some individuals remain illiterate, nor do we throw away our automobiles and radios because they fail to function properly at times. Let us face the issue squarely. What we need is more basic research and less exploitation.

Large universities and technical schools should be encouraged to conduct fundamental researches of joint interest to psychologists and engineers. It seems reasonable that studies now in progress should be used as a point of departure. Much time and energy have already been lost through duplication. As technical schools become genuinely interested and show initiative of their own, the work may be broadened. Interest should precede aid given. Sporadic interest will be engendered in any field if large enough foundations are established. The real test comes when the foundation ceases. Do those in charge intelligently continue the studies in which they profess an interest?

One of the urgent needs now is administrative and financial provision for scientific research on the problems of automo-

bile driving. Through foresight some administrators have sensed the need for such studies and are allowing time and support for staff members in this field. Others have been very slow to see the need for work along this line. At present, there is enough genuine interest to get things going if administrative set-ups are made conducive to productive results.

By nature of the problems involved, funds are necessary. It is easy to mimeograph a questionnaire and mail it to 1000 addresses and wait for a ten or twenty per cent return which can be tabulated and analyzed at some length. It is also easy to "rig up" laboratory apparatus to measure the contraction of the pupil of the eye under different degrees of illumination. It is quite another matter to standardize and evaluate two tons of equipment which must be taken by truck into commercial plants for standardization and validation studies, using proper subjects for such purposes—the driver who is working on the job. A population of 100, 1000 or even 10,000 sophomore students, having a literary bent, hardly constitutes a representative population for evaluating automobile driving. Perhaps our intelligence tests would have been in better repute today if the earlier workers had departed from a purely academic criterion.

It is hoped that this paper may cause some serious thought on one of the greatest national problems facing our country today, a calamity which is all too often taken for granted. It is time for psychologists to come off the sidelines and get into the game if our science is to do the service to society which is expected of it. An academic attitude is usually quite safe—safe from criticism and progress—but it will not help the world solve its problems. Physics, chemistry, the biological sciences and mathematics are fast taking their place in the field of applied science. Genetics and many newer sciences such as bacteriology are being used to make the world a more desirable place for man to live. And yet we have a few "crepe-hangers" who will wisely ask, "How do you know our world is better than formerly?"

Perhaps it may take another generation to lead us out of the muddle in which the "semi-philosopher-scientist" finds himself. A half-caste, he is too scientific to speculate in true philosophical fashion, yet too philosophical to get down to business and work up a first-class "experimental sweat" to properly orient himself.

Let us not deceive ourselves. The problem of highway and traffic safety is going to require many arduous hours of "hammer-and-saw" effort to assure fruitful results. Those averse to work and discouragement should retire early. And yet there must be a solution. In the meantime, we must be on the sharp lookout for blind alleys. If perchance we follow one to its termination, there are plenty of avenues remaining to be explored. One or more of them may lead us where we wish to go. The philosophical implications here need not be considered very profound. The fundamental objective is conservation of human life. It is a problem which cannot be considered lightly.

VIEWING THE TRAFFIC PROBLEM

HERBERT A. TOOPS
Ohio State University

S. EDSON HAVEN
Wayne University

INTRODUCTION

IN reading the news clippings and articles on traffic accidents submitted by our students, we became so disgusted with the inane nature and lack of scientific character of some of the pronouncements from the press, radio and platform that we decided to make an investigation; even though we had never done any previous experimentation or research in the field or been more than mildly interested therein.

Our justification for such unseemly behavior derives from the well supported conclusion that the faults of a school system are more readily noted by an outside surveyor than by members of a local investigating board. Here, the point of view of the stranger to the fold has been found precious. He sees details in their perspective; his eyes fail not to see the forest for the trees. He looks for principles, and takes difficulty of execution for granted. He dubs nothing precious simply because it cost much to attain. He sees faults because of their novelty; their obviousness has not made him blind. Criticism from outside the profession, then, has much of potential merit, a deduction not inconsistent with the well-authenticated observation that such stranger-determined faults, more readily than the self-discovered, are allowed to accumulate dust on the top shelves of a much unused storage vault.

In our endeavor, we resolved to use our eyes in observing what might be seen, and our scientific judgment in evaluating what might be read or heard about the matter. We found

faults innumerable, but are quite willing to leave to others their correction.

These defects, in our opinion, any citizen conceivably might see if only education were not so severely condemnatory of, and the executive offices of mankind so universally inhospitable to, criticism as a means of social amelioration. When every teacher has the attitude: "Who art thou, worm, to criticize; greater men than thee have built what thou dost see," what surprise that the adult who becomes ameliorizer should be deeply resentful of all criticism save that of equals and superiors in his own profession, and, by implication should indict all men of other professions as fools should they dare to tread upon one's sacred domain!

In the field of accidents, we accept from the public press grisly accounts which frighten us into ineptitude, and pay taxes for the support of safety programs which any scientist of human nature might tell off-hand have not the slightest chance of success. Eyes that see not and ears that hear not seem to characterize him who writes on the subject. Yet we find that drivers of cars—without which element crashes are all but impossible—are more eager to talk about their accidents and near-accidents than almost any other topic, not excepting the weather, health and surgical operations.

SOURCES OF INFORMATION

In order that we might gain some insight into the various phases of the problem, we made a systematic examination of several types of available evidence. These important sources of data were as follows:

1. The official statistical reports of auto accidents of the forty-eight states, only slightly more than half of which have regularly organized traffic or record bureaus.
2. Clippings collected, personally, from metropolitan dailies, and from magazines; also, clippings from similar sources by correspondents residing in several large cities. And others, handed in by university students, collected, in the main, from their home-town papers.

3. "Accident Prevention" literature from insurance and indemnity companies.

4. Some of the more scientific publications on the subject by psychologists, engineers, physiologists, optometrists, psychiatrists, and illumination experts.

4. Examination of parked autos for residual evidences of accidents.

6. Stenotyped records of traffic court proceedings broadcast over the radio.

7. Interviews with drivers carried out according to a formal plan by psychology students.

8. Personally conducted interviews with bus, taxi, truck, and other drivers.

9. Direct observation of traffic court proceedings.

10. Visits to homes of traffic victims.

11. Interviews with service station attendants and garagemen.

12. Interviews with pedestrians who do not drive.

The interviews with hundreds of motorists were our most productive source of information. These drivers had escaped the Demon Monster with little more damage than the occasional loss of an arm or leg, even though in many cases, their companions or the occupants of "the other car" had not come off so lucky. Friend will reveal to Friend what Citizen-at-large, and especially what Citizen-in-jeopardy, will not reveal to Policeman, or to Policeman-in-any-other-garb. Moreover, our interviewers had at least a modicum of training in the observation of human behavior, and to that extent were able to sift aspiration from observation, rationalization from motive, and inference from fact. Further, the method was inexpensive, swift and, we believe, effective.

In collecting specific experiences in both normal driving and in accidents, the interviewers reported, among other items, upon the driver's answers to the following questions: (a) What mishaps have you ever had with a car? (b) Where were you going? (c) Was your accident reported? (d) Have you ever had any narrow escapes? (e) What are you most afraid of in driving? (f) What about other drivers peeves you most? (g) When are you most dangerous as a

driver? (h) Do you have any rules that you follow in driving? (i) What are some of the legitimate occasions for speed? That is, under what circumstances is "speeding" permissible?

The details of our utilization of the other methods, and even of this, are of little importance in other than a strictly technical paper; therefore, suffice it here to say that we analyzed the reports, tabulated the answers, and examined carefully the accident statistics.

EVALUATION OF EVIDENCE

As anticipated, much of the literature and oral testimony revealed a total ignorance of simple correlational analysis, not to mention such erudite procedures as partial correlations and multiple regression equations, with the expected muddiness of thinking—to dub the thinking by the least opprobrious term which seems charitable under the circumstances.

The official, published accident statistics give frequency tabulations only; no two-way charts are used, due in part to the fact that no adequate notion of the criterion, as the psychologists know it, is ever employed. Further, the statistics are not reported in standard terminology. And, worse, the whole accident study approach is highly seasoned with moralistic conceptions. As witness of this, note how one puritanical state designates its chief causes of auto accidents: "Inattentive use of the highway; lack of care approaching pedestrian in street or obstruction to view; inattention to driving or traffic; too fast for course or road surface, traffic, weather, or locality; failure to slow down at intersecting street; improper turning; careless backing; improper use of highways, alcoholic operation; defective passing." And four others, only two of which were objective, however.

The person who drives an automobile for any considerable time without an accident is a rare individual. Observe in the fenders of the parked cars as you go down the street the unmistakable evidence of the continual warfare of fender upon fender and steel against steel, and you will be truly amazed.

In one group of two hundred interview reports from drivers, one hundred sixty-six had been involved in accidents (only one admitted that he was at fault). Of course, there may have been some selection here, although none was intended, and no deliberate effort was made to secure randomness of sampling.

A smaller proportion of accidents are fatal than the statistics indicate. Many are never reported to the police or traffic bureau; in fact, a majority are not. The fatal ones, however, are reported with some fidelity, although there are several states with no central automobile accident statistical bureau. Non-fatal mishaps are of as great, or even greater, import in understanding the psychological causes as the fatal, because, obviously, "Dead men tell no tales." Since society does not approve of rigidly controlled experiments in which accident hazards are introduced or even manipulated with a view to determining accident potentialities of various factors, we need to find out all we can from those who have been "all but killed." The dilemma of young Doctor Arrowsmith besets the state highway engineer upon obtaining an idea of traffic improvement which looks "good."

States do not use a common definition of accidents or a common form in reporting them; nor are there used in the majority of instances objectively verifiable categories such that a scientist might use them for analysis into the possible causes upon which a sound accident prevention program might be based. Following are some of the worst examples appearing on the official forms: "Occupant improperly located; operating a defective motor vehicle; failure to grant the right-of-way; using the street inattentively; exceeding the speed limit; bad physical defect; under the influence of alcohol; glaring head-lights."

JOURNALISTIC TRADITIONS

For the most part the public gets its ideas regarding accidents and accident-prevention from the newspapers, occasion-

ally from magazines; and more recently from the screen and radio. As in all public opinion, the attitude which results depends on what stimulating situation is presented.

The press has been very influential in formulating public ideas and unfortunately has perpetuated a number of mental stereotypes; and at the moment is busily engaged in the promulgation of many other beliefs of equally questionable validity. The public has been taught that Speed, Gin and Recklessness are everywhere the wicked triumvirate which produces the grim spectre of death on the highway; and, Engineering, Education and Rigid Enforcement are the way out.

All of our evidence indicates that speed—meaning fast driving—is far from being the most important variable. Much to our surprise low speed (fifteen miles per hour, perhaps) is among the most dreaded menaces of the highway, if we may trust the judgment of our interviewees.

In accordance with the view that high speed is an important cause of crashes the press has often advocated a governor which would physically restrain cars to a certain upper limit. Rather than a mechanical device to prevent excess speed, the crying need, so our interviews with good drivers indicate, is for a governor-actuated instrument to emit a characteristic warning signal when the retreating vehicle falls below a given speed thus adding to the accident hazard by imposing unnecessary interference to traffic moving in the same direction. (This raises the question: How shall we equip an automobile so that when driven safely it shall be normally colored, but when unsafe through careless driving it shall turn *red?*)

Even with all law enforcement officers on the lookout to detect drivers soaked with liquor there are few convictions for driving while intoxicated. It is clear that alcoholic beverages are no greater cause in filling the jails with reckless drivers than their absence, after 1919, was in emptying the jails, almshouses, and asylums, as promised. We believe, however, that the emphasis of the press upon the danger of driving while "drunk" is justified.

Recklessness is not a useful psychological category. The driver from his own point of view is never reckless, and from the psychological standpoint, what we call "recklessness" is better measured by such statistical variables as speed, driving, skill, attitudinal habits, familiarity with the road, and the like, all of the "incommensurables" of which can be measured by tests and techniques for the most part yet to be developed. Motorists tell us that taxi drivers are highly skilled, not reckless, but that if the average citizen attempted to imitate the public chauffer he would indeed be reckless. Accordingly, for our purposes, we may conclude that "recklessness" is a moral judgment, rather than a scientific datum.

We may observe further trends with respect to published statements:

1. A tendency to exploit the horrors of the accident itself; and dropping it at this point, rather than following it to a logical conclusion of revealing to the readers the anguish produced in the life of the maimed and his family, by reason of the accident.
2. Reticence in mentioning the make of car involved. Surely, there are wide individual differences in the accident potentialities of automobiles.
3. Willingness to sponsor restrictive measures on the autoist's freedom.
4. Willingness to sponsor frequent traffic crusades, which are of more value to a depleted public treasury than a permanent solution of the traffic problem.
5. To play up sex differences among drivers, however insignificant or statistically undependable.
6. To make a great fuss over every (small) reduction in the accident rate. Every statistician appreciates that small reductions in the rate are almost certainly ascribable to mere chance and so are no evidence of a trend.

The account does not stop here. Additional views in process of formulation are:

1. That blowouts are exceptionally hazardous. The statistics show they are relatively infrequent. And, in our opinion, a blowout may be the *result* where it is assumed to be the *cause*.

2. That individual reaction time is the all-important measure of the individual's ability to drive.

3. That increased illumination will solve the perils of night driving. (Our only comment is that Old Sol fails to do it in the daytime!)

4. That rigorous law enforcement will solve the problem. Any utterance for greater severity of punishment or stricter literal adherence to the law is usually highly applauded by platform and bench, and is copiously reported in the news.

5. That by naming an unknown cause, the accident is explained: for example, "Defective brakes," and "Intoxicated driver." On a par, in our estimation, with "Defective wiring" and "Defective flues" as causes of the unexplainable destruction of houses by fire.

Any classification, such as the above, is bound to do some injustice to the general situation which it purports to summarize. However, it is obvious that the positive contributions of the public press may be minimized by reason of the prevailing traditions.

BUNGLING PREVENTION

Under pressure of press, bench and platform, the public has become "punishment-minded" rather than "preventive-minded." In minor cases of traffic law infringement if the culprit does not elect at the instigation of officers to purchase his freedom as he would so much butter, or to secure the same results through the aid of fixer friends, mostly political, he comes before the court to be judged rather than to be tried. The judge who should function as the educational agent of the public does not have available the cumulative record of past experiences, individual tendencies, and measures of skill and aptitudes which characterize the progressive work of the juvenile court. The personnel philosophy which enjoins the weighing of the individual's reactions in the light of his personal motives and capacities is conspicuous by its absence in the treatment of the traffic offender. Any personnel officer in a college without cumulative records would demand the immediate setting-up of such a system in order that he should not make his judgments regarding students' welfare in ignorance of the ascertainable facts.

DANGEROUS PEDESTRIANS

Our talks with pedestrians reveal that they have a child-like and mistaken faith in the ability of autoists to drive safely if they only would. It is unreasonable to expect a one-and-three-quarters ton projectile at forty miles per hour to be as adaptable as a 160-pound projectile traveling at two-and-a half miles per hour. The former is a car; the latter a man. The smashing power of the auto under the above conditions is five thousand times that of the man. Accordingly, in a collision, the man crumples like a grasshopper. This is one reason for the high fatality rate of pedestrian-auto collisions. Is not the conclusion evident, then, that some of the legislation designed to restrict the auto should be applied as well to the pedestrian? In view of the danger is it unreasonable to expect every pedestrian to be compelled to wear reflecting buttons, white overshoes, or to dress in white clothing, and walk facing traffic under a legal presumption of "reckless walking" in case of accident?

"ACCEPTABLE CAUSES"

One of the gravest defects of accident prevention work is the failure of Traffic Bureaus, both state and city, to employ scientists capable of constructing questionnaire forms and record blanks to obtain the necessary data for correlational analysis. In other words, a determination of the factors which are related to safe operation. Many variables worthy of being collected are not now included in the reports; for example, height and weight of driver; occupation; "Index of foreignness," that is, degree of familiarity of the driver with the road where the accident occurred; number of occupants in the car; make and model of car; temperature; driving experience; driver's intent, where going; and, number of dependents and their ages.

Accident categories must be couched in terms of verifiable human attitudes, skills, and habits; and, objective classifications regarding cars and roads. Then, with dependable and trustworthy primary data, it will be possible to interpret the

ascertained facts by a refined statistics which will dig beneath the surface manifestations for the "basic causes."

SAFER CARS

Our data suggest that there is a great need for improvement of cars. For example, in accidents, door-handles have been shown to rip open people's heads and bodies, suggesting their being recessed into the door. The running-board is a hazard to spirited youth; although the street car long since adopted the retractable step. The lights on all cars should be elevated for greater effectiveness, and if countersunk near the top, the outline of the "monster" would be more easily discerned. The front corner-posts should be reduced to as near zero in size as possible, and the rear should be a replica of the front in this respect. Ships of the air, and of the sea, and submarines beneath its surface, all carry a mast or periscope to ward off or to detect danger. Only the ships of the land lack these warning beacons.

BETTER HIGHWAYS

There is great need for improvement in traffic control devices, road construction and touring conveniences. It is obvious that we should seek to remove hazards by engineering applications rather than attempt to adjust people to a defective environment.

We have turn-signs, for example, only because there is a bend in the road. Such a "distractor" is there because the road is defective; because, in short, the drivers who approach it, must do something which they would not have to do if the road were better constructed. The sign is there because the average motorist must slow down, and a definite twist of the wheel must be executed in order to prevent a collision with oncoming vehicles, fences and any other impediments. This suggests as a statistical measure of the goodness of a completed highway the standard deviation of the speeds of travel on it. The good road is one upon which uniform speed is not only possible but actually prevails. (If we had a continuous

record of the driver's speed, a low index of variability of the speeds would probably be a fair measure of the carefulness of the driver.)

For some time in the future, however, we shall have need for signs to overcome the defects of our streets and highways. Therefore, as a general principle we should give the driver the specifications of the danger and grant him the judgment as to the reactions appropriate to the circumstances. In other words, we should describe the *situation* and leave to him such decisions as, "Go Down In Second Gear." People, generally, do not like the impartiality of those mechanical devices which replace intelligence.

Instead of saying, "Dangerous Hill" or "Hill—Reduce Speed," which are judgments on the part of someone who does not know *our* automobile, we might say, "Hill—5°" or "Hill —10°." After having experienced a few ten-degree vertical curves, correctly labeled, we would know what is to be done with *our* car hereafter on ten-degree hills! Ten-degree hills may be bad, very bad, or dangerous, and "require first gear," depending on our machine, our driving habits, the weather, time of day, and so on.

Cross-roads should be indicated in the same manner. Suppose a large sign tells us, "Cross Traffic—50 Miles Per Hour"; we would sit up and take notice and treat the situation for the risk involved.

There are many useful as well as useless signs. To all of these the driver, perforce, must pay attention for fear of not noting something essential. For example, one occasionally observes on a main-traveled thoroughfare such signs as: "Tractors With Lugs Prohibited." Now ten thousand persons per day, say, needs must look at that warning, while only the three hundred thousandth would have any temptation to drive a "tractor with lugs" on that pavement.

In any survey of a locality from the viewpoint of accident hazards, the effectiveness of the signs might be tallied up on a score-card both for daylight and night driving by a total stranger to the community. He can see the faults to which

the people at home through long usage have become blinded. The community acts on its knowledge of the situation, and its knowledge may be misinformation. The stranger acts on what he thinks he observes.

Our interviewees tell us that there is need for "rest zones" along major routes where the drowsy or over-tired individual who is aware of his driving untrustworthiness may stop to relax in safety. A fifteen-minute rest pause may restore almost normal proficiency for several hours of additional driving. The servicing of the driver and occupants in addition to servicing the car is only imperfectly realized by oil companies.

"Zones of reflection," out of the line of traffic on all main roads, located at Y's and important intersections, where judgments for future course of action are involved, such as examining maps, would be an adjunct to safety. Turn-out zones in front of farm houses are necessary for school buses, mail carriers, and other deliverymen.

MOTIVATING SAFETY

We may effect many improvements in driving for which the driver himself is responsible. (a) By publicly encouraging him to refuse to get behind the wheel when he feels himself incompetent. The papers recently reported that in one city, upon call, a policeman would come to drive home anyone who felt himself unfit to operate his car. We might wish such official attitudes were more frequent! Our interviewees feel that they are most dangerous when they are in a hurry, especially going to and from work; sleepy; fatigued from continuous driving; and intoxicated. (b) By selecting a driving period when hazards are at a minimum; traffic density less. (c) By affording means for the driver to criticise the weakness of his fellow-men through the use of suggestion boxes placed in all service centers and first aid stations by the state highway department. Criticisms accruing to the license number would be used in a purely clinical fashion. (d) Compulsory examination on a valid diagnostic test of all persons making their

first application for a driving license, and after every accident reported to the police, with clinical treatment, compulsory, if necessary.

It would seem the better psychology, in accident-prevention programs, to emphasize proper driving habits; to encourage the periodic self-inspection of cars and of driving skills; to use reward rather than punishment or penalties; and, to soft-pedal gruesomeness and tragedy. A driver does not gain competence or develop mastery under the stress of fear and the conviction of error. It is apparent that safety needs to be sold to the public according to the same motivational recipe as any other commodity or service—for example, education.

CONCLUSION

The day of panaceas in handling the complex phases of the motor traffic problem is past. We are rapidly recognizing that an adequate solution will require the cooperative effort and unselfish service of many professions:—the auto designer, road builder, statistician, psychologist, legislator, judge, sociologist, health officer, publicist, and city planner, and the to-be-developed expert in national planning.

REFERENCES

1. American Standards Association. "Manual on Uniform Traffic Control Devices for Streets and Highways." Washington, D. C., 1935.
2. Annual Statistical Reports of Various Motor Vehicle Departments and Traffic Bureaus for 1934 and 1935.
3. DeSilva, Harry R. "Research on Driving Skill." Boston, 1935.
4. Lauer, Alvhh R. "Methods of Measuring the Ability to Drive an Automobile." Iowa State College, Bulletin 115, 1936.
5. National Safety Council. "Accident Facts." Chicago, 1935.
6. Proceedings of the Highway Research Board. Washington, D. C., 1933, 1934, and 1935.
7. Stoeckel, R. B., May, M. A., and Kirby, R. S. "Sense and Safety on the Road." D. Appleton-Century Co., 1936.
8. Travelers Insurance Company. "Thou Shalt Not Kill." Hartford, Conn., 1935.
9. Weiss, A. P., and Lauer, A. R. "Psychological Principles of Automobile Driving." Ohio State University, Contributions in Psychology No. 11, 1931.

THE TACTILE–KINAESTHETIC PERCEPTION OF FABRICS WITH EMPHASIS ON THEIR RELATIVE PLEASANTNESS

ROWENA RIPIN
Psychological Corporation
AND
PAUL F. LAZARSFELD
University of Newark Research Center

NOTE: The following pages are part of a study which was initiated and continuously advised by Alexis Sommoripa of the DuPont Company. His comments are to be found at the end of this report.—H. C. LINK.

I. INTRODUCTORY

THE following is based on a study undertaken by the Psychological Corporation at the instigation and expense of a commercial concern of rayon manufacturers. These sponsors were interested in learning what purely tactile-kinaesthetic distinctions could be made between fabrics in general and specifically, between silks and rayons, and how these distinguishable attributes influenced preferences. They were interested only in the response of women, representative of the fabric buying public. Weight was included along with the tactile attributes inasmuch as it plays an important part in the judgment of fabrics and could be perceived by blindfolded subjects. The results discussed in this paper are selected on the basis of their psychological interest. Those referring more specifically to the fabrics themselves are omitted.

The usual study in "hedonic psychophysics"[1] employs a series of stimuli differing systematically with respect to one or more known variables. This procedure could not

[1] Term from Beebe-Center (1).

be employed in the present study for no such qualitative categories have been set up for the sense of touch so far as the experimenters know. The range of tactile stimulation may be so wide as to defy such a classification or the study of the field may not yet have reached this stage of knowledge. It was necessary, therefore, for the experimenters to undertake a preliminary, explorative investigation on a few observers (O's) before proceeding to work with a large number in order that they might learn to what extent the sensory elements involved could be systematized and in how far this would be necessary for uncovering the hedonic factors.

In addition to this preliminary study in which 48 fabrics were presented to each of 6 O's, the experimental procedure proper consisted of four series. It lasted about one hour with each O. The O's all came to the room where the experiment was set up. In so far as the study dealt with touch responses, visual impressions were excluded by blindfolding the O's. They were seated in a chair before a rod over which the various fabrics, all cut to the same size, were hung.

The observers were a hundred women, distributed as equally as possible among the following (cross) categories: college and non-college; low, middle, and high budget groups; and these age groups: below 20, 20–25, 25–30, 30–35, 35–45, and over 45.

All the fabrics used were chosen in collaboration with textile authorities.

This paper deals with only one of the experimental series and its results, a detailed description of which follows.

Comparison of Pairs

The *purpose* of this series was to measure the relative preferences for silks and rayons and to learn the reasons for such preferences.

The *materials* consisted of ten pairs of fine, expensive fabrics with a silk and rayon member for each pair, matched as carefully as possible in all attributes other than this one of fibre. The method for choosing members of a pair was the following:

On the basis of the preliminary series, the experimenters set up groups of five fabrics which they deemed objectively similar. Five blindfolded O's were asked which two out of the five fabrics were most similar. The two most often paired in this way were chosen for the final pair. In order of presentation the pairs were: medium crepe, velvet, triple sheer, satin, heavy crepe, taffeta, crepe, heavy sheer, flat print, and a fourth crepe.

The *procedure* was as follows: Our preliminary observations had shown that O's had difficulty in concentrating on their likes and dislikes in the motives for their preferences and that they would readily slip back into a mere description of the fabrics. In order to enforce the hedonic, evaluating attitude which we were interested in arousing, we inserted the procedure of rating the fabrics for pleasantness prior to the O's statement of preference. We had also found that so simple and direct an approach as asking an O why she liked a particular fabric or attitude would result, in the great majority of cases, in the O being at a loss to answer. Our O's were therefore presented with two fabrics and were asked to tell why they *preferred* one to the other. Such paired comparisons sharpened their attitude of observation, limited the scope of their comments, and provided pressure and purpose for their aesthetic processes.

Preliminary Procedure

In order to familiarize the O with the set-up, a preliminary presentation was made during which instructions were given for the rating scale to be used. The O was told to feel the fabric and give it a rating from one to five, such that "one" would represent indifference, "five" liking the fabric very much, and "two," "three," and "four," the degrees of liking in between. If the O disliked the fabric she was to say so.* After rating, the O's blindfold was removed and she was asked to feel and rate the fabric again while looking at it. The experimenter took this opportunity of demonstrating the

* In view of the selection of fabrics this practically never occurred.

various movements of touch to be brought into use by the O; feeling the surface between two fingers, the long up and down sweeping movements, crushing, and weighing. If the O's rating changed after seeing the fabric, she was asked why.

Pleasantness Ratings

One member of the pair was presented to the O who was instructed to feel of it and give it a rating. The second member was next presented with the same procedure. Silks and rayons were alternated as first members from pair to pair.

Preferential Choices

Then both fabrics were presented together and the O was asked for her preference. If an O could not choose between the two, she was not urged to; if she gave a choice, however, she was urged to give her reasons for it. All of the O's remarks and any characteristic activity were entered on the record sheet.

In the second series, six triplets of fabrics were presented, each consisting of fabrics representing three degrees of some one varying attribute. The O's were asked to rank the triplet members with respect to the variable and to state which one they preferred. In the third series, seven fabrics were presented and the O was asked to make a first, second, third, and last choice among them for a dress. In the fourth, the same seven fabrics were presented and the O was instructed to eliminate one fabric at a time, choosing the one which seemed most unlike all the others. In all cases the O's were blindfolded.

II. CLASSIFICATION OF TACTILE EXPERIENCES

Rayon-Silk Preferences

One of the most important results from the point of view of the sponsors of the study was the relative popularity of the rayon and the silk fabrics. As this is of no psychological interest, we shall state it here in only the following summary figures:

*Average no. of choices favoring rayon	53.6
Average no. of choices favoring silk	30.9
Average no. of no choices	15.5

* The average for all ten pairs.

When the average number of choices for rayon was corrected so as to include the proper proportional amount of the number of no choices, it amounted to 64, which is to say, that for every hundred choices made, 64 were in favor of rayon and 36 were in favor of silk.

Analysis According to Content

Each comment made was recorded, as was the number of times it appeared in connection with each fabric. Since all comments were given as reasons for a preference, they are all positive with respect to feeling tone. This is not only an unwieldy body of data as it stands in itself but it is also meaningless. To quote their totals: 285 for *smooth*, 154 for *fine*, 33 for *rough*, 354 for *soft*, etc., could lead to no significant results.

The first difficulty in the way of interpreting and utilizing these data is, of course, the question as to the accuracy and reliability of the O's reports. We do not know, *e.g.*, whether the experience on one occasion reported as "fine" might on another be termed "silky." "Silky" and "fine" are terms representing different shades of experience but they share a significant element of meaning in common. In order to extract and emphasize this common factor, and cancel out the minor discrepancies, it is advisable to lump individual nuances into a concept larger than a collective term. This procedure is further justified by our evidence that a single attribute is rarely experienced as such and is rarely judged aesthetically by itself. (See below.) For meaningful conclusions, we should refer, therefore, to appropriate combinations of attributes.

Proceeding along these lines of large concepts and configurations of attributes, we present a distinction which gradually became apparent from our O's comments and which we were

able to check from several statistical approaches. It is a distinction between two groups of subjective attributes which we have termed here the *demanding* and *relaxing*. This appelation is based on the consideration that "relaxing" attributes require the least expenditure of energy on the part of the O for their reception, whereas "demanding" by right of the physical resistance they offer in the way of friction, or to pressure, etc., require a greater expenditure of energy for their reception. These terms, hereafter referred to as D and X for "demanding" and "relaxing" respectively, will become clear and can best be defined by listing the terms subsumed under each. For this classification all comments with the exception of the kinaesthetic have been included. Along with the tactile it includes terms descriptive of feeling tones and "character," as we are not so much concerned here with classifying textile attributes but with differentiation between two types of pleasantness in tactile experiences.

We understand the term "relaxing" to apply to the following quoted comments:

smooth, silky
fine, finer quality
even, threads don't stand out so much, even in the nap, closer texture
soft, pliable, melts in your hands, clings more (which suggests that I should think it would be finer)
sheer, thin
more appealing, more pleasing, more comfortable
goes with delicacy, caress, frivolous, gossamer, airy, luxurious, gentler

We understand the term "demanding" to apply to the following:

rougher, crepier, not quite so satiny, not quite so slippery, doesn't slip, a little more woolish, not so shiny
strong grain, looser weave, more firmly woven, coarser weave
stiffer, spring to it, more crisp, more resistance
more elastic

firmer, stronger
thicker, more substantial, more body
wear longer
more noise
more character, not neither-nor, more interesting
doesn't try to be something it isn't
don't like silks

Frequency of H–L and X–D Attributes

The first justification for the X–D classification lies in the fact that it was possible to attribute to most of the reasons for preference either a relaxing or a demanding attribute. Only in 5.7 per cent of all the choices were comments made which forced us to register the choice as one implying both relaxing and demanding experiences. Comments for weight were held separately. We can, therefore, list the reasons for preference under the following heads: the relaxing choices (X), the demanding choices (D), their combination (XD); likewise the heavy (H) and the light (L) weight choices. These are found in Table I.

TABLE I

*Quantitative Distribution of Nature of Preferences**

X (Relaxing)	600
D (Demanding)	181
H (Heavy)	127
L (Light)	158

* In 15.5% of the cases there was no choice and hence no comment, this table covers the remainder of the cases; there is overlapping in that one choice could depend upon both H-L and X-D comments, and where the choice was of an XD nature, it was counted twice, once as X and once as D.

From these figures, it is obvious that relaxing experiences are by far the most frequent reasons for fabric preferences, these being more than four times as numerous as the demanding.

H–L and X–D Preferences and Rayon-Silk Differentiation

The question arises whether the nature of the preferences bears any relation to the quality of the fibre of the fabric.

To take the weight preferences first, we may state that the actual number of times weight was cited as a reason for preference was relatively the same for rayons and silks, 160 and 125 respectively. When the weight comments are divided into heavy and light, however, it becomes apparent that silks are chosen relatively more often because of their lightness, rayons because of their heaviness. This distribution is expressed in percentages in Table 2. The difference is quite striking numerically but from our limited data it is impossible to tell whether this is a true difference intrinsically determined by the nature of the two fibres or merely a difference determined by the sampling of these fibres as used in our study.

TABLE II

Distribution of H–L Preferences for Rayons and Silks

	RAYONS	SILKS	BOTH
Heavy	62%	22%	45%
Light	38%	78%	55%
Total	100%	100%	100%

The distribution of X–D preferences according to rayons and silks is given in percentages in Table III.

TABLE III

Distribution of X–D Preferences for Rayons and Silks

	RAYONS	SILKS	BOTH
X (Relaxing)	86%	74%	81%
D (Demanding)	14%	26%	19%
Total	100%	100%	100%

For both rayons and silks the relaxing attributes are given far more frequently as reasons for preference, but, relatively, these occur slightly more frequently for rayons than for silks.

These figures represent, of course, the *average* for all pairs. Much more distinct but antagonistic differences between the rayons and silks are brought out in selected pairs. In pairs I, X, and IX, the medium crepes, light crepes, and flats, respectively, the prevalence of relaxing reasons for the rayons are as distinct as that of demanding for the silks:

	RAYONS	SILKS	BOTH
X	97%	33%	84%
D	3%	67%	16%
Total	100%	100%	100%

Three other pairs—II, VI, and V, the velvets, taffetas, and heavy crepes respectively—yield the reverse distributions. Here the vast majority of relaxing sensations are attributed to the silks:

	RAYONS	SILKS	BOTH
X	60%	93%	80%
D	40%	7%	30%
Total	100%	100%	100%

X–D Preferences and Fabric Types

The X–D preference distribution may be considered by comparing the relative amount of X and D reasons given in the choices for each fabric pair. This is most easily expressed in terms of a ratio which would indicate the number of D choices for every hundred X choices made for each pair as shown in Table IV.

TABLE IV

D/X Ratios (The Number of D Choices for Every 100 X Choices)

PAIR NUMBER	FABRIC	RATIO
IX	Flat prints	14.0
II	Velvets	14.2
VII	Medium crepes	17.5
I	Medium crepes	22.2
IV	Satins	23.6
X	Light crepes	25.2
II	Triple sheers	27.8
VI	Taffetas	37.6
V	Heavy crepes	41.0
VIII	Heavy crepes	44.5

A consideration of these figures together with that of the objective features of the fabrics involved leads us to the following tentative statement. There are some indications that

predilection for demanding attributes increases with the objective degree of non-soft and non-smooth features of the fabrics under comparison. Referring to the above table of D/X ratios we find that the velvets and flat prints yield the lowest index, 14. Objectively these fabrics are probably the softest and smoothest of all those used in the study. At the other end, the heavy sheers and the heavy crepes yield the relatively high indexes, 44.5 and 41. According to this hypothesis, then, an objective increase in the demanding quality in two fabrics brings about a relative increase in the number of demanding attributes given for preference. (See Aesthetic Considerations below).

Distribution of X–D Attributes According to Qualitative Dimensions

In the face of great difficulty in classifying the tactile comments as such, we used the data from the Preliminary Series to set up the following two-fold division in which the touch activity of the O presents criteria as well as the descriptive, tactile comments reported.

Surface attributes: Those attributes which are perceived by moving the hand along the surface with a sliding movement or by feeling the fabric between the fingers. This includes the larger sub-heads of *rough* and *smooth:* weave and thread distinctions which may be detected in this way, chiefly *fine* and *coarse;* surface temperature; and comments on the *nap.*

Three-dimensional attributes of depth as over against surface: those attributes which are perceived by crushing and folding the material (*i.e.*, gathering into folds). Its subheads are the various types of *resistance* the material can offer, such as *softness* and *stiffness; firmness, strength, elasticity, clinging character; thickness, body,* and their opposites.

The two-fold classification into surface and depth dimensions has its emphasis on the sensory qualities involved and hence tends toward the objective. The distinction between relaxing and demanding attributes has its emphasis on respon-

sive qualities and hence tends toward the subjective.[1] It is interesting to note how these qualities are distributed with reference to one another. Table V gives the figures for the

TABLE V

Number of X and D Comments for Surface and Depth Dimensions

QUALITATIVE DIMENSION	X	D	TOTAL
Surface	439	44	483
Depth	370	121	491
Total	809	165	974

number of comments of relaxing and demanding nature made with respect to surface and depth characteristics.

It is apparent that the demanding comments refer chiefly to depth characteristics: resilient, firm, etc. Pleasantness associated with demanding attributes in the surface category is rare; there is almost no enjoyment reported for roughness.[2]

III. TACTILE EXPERIENCES IN THEIR MUTUAL INTERRELATIONS

One of the most interesting results of the study was the limitation to analysis soon encountered when dealing with the tactile experiences.* The question of our starting point was, *just what* about a fabric is liked and *just what* is disliked? The answer to questions of this nature should be in terms of elements but we were unable to determine these as such, let alone set up a hedonic hierarchy among them.

We found that judgments of elements were so influenced by the totality of experience in which they occurred that a consideration of the one without the other was meaningless. We

[1] Our use of the terms—objective and subjective—does not coincide with Katz's. (2) According to him, tactile perceptions are objective when localized in the environment; they are subjective when localized within the 0. In this sense, all our categories are objective.

[2] Major found stiffness, roughness, and coarseness to be unpleasant; softness and smoothness, pleasant; fineness, thickness, and thinness, variable (3).

* This section utilizes the material from the preliminary experiments in addition to that from the paired comparisons.

can divide the data with respect to this point into two groups: the first where the larger frame of reference for the judgment is the configuration of what is sensorially given; the second where the larger frame of reference includes the O's related experiences, *i.e.*, cognitive factors in addition to those directly given by the stimulus.

The Influence of the Tactile Configuration

To begin with the simplest case of purely tactile judgments, we should expect the most elementary statements to be of the order "this is rough," or "this is smooth." Such remarks were made by some O's, though rarely, and usually where there was a lack of vocabulary to the extent that the O, herself, apologized for it. So far as we can judge, our O's did not as a general thing receive an impression of roughness *per se*, but immediately perceived the roughness as being one of fibre or of a recurring pattern in the weave, or as of crepiness, or as of imperfections in the fabric's finish. In other words, their experiences were not reported as varying degrees of a tactile quality but as *characteristics of a fabric*. This was brought out even more clearly when the O's blindfold was removed (Preliminary Series I) and visual data added their influence to that of the total tactile-kinaesthetic configuration. Although the objective roughness of the fabric was the same, roughness judgments were changed as follows:

"Not so rough this way"
"More definite roughness this way"
"When I didn't see the material I felt too many knots"

Often O's will mention two or more attributes in combination. It is not only that attributes are paired in the O's reports, but that the attribute seems to be experienced as a function of the combination in which it occurs. Occasionally, two members of such a combination will become confused, one with the other, as in the following examples for thickness and weight.

"The heaviness isn't the thickness, it's the weight"
"Quite heavy—not the weight, but the thickness"
"Fairly light and graceful though not actually light in weight"

The following quotations illustrate the strong interaction in combinations of weight with extreme smoothness.

"It gives the impression of being heavy because the surface is not so smooth"
"Fabric $15 is light . . . no, maybe in actual weight $12 is lighter, but I suppose $15 is so smooth it feels lighter"
"Feels either rougher or heavier"

The same influence is at work in preferential judgments, possibly to a greater extent. Often O's expressed themselves as taking a general stand toward a certain attribute: "I like roughness" or "I favor soft materials," etc. But even O's with expressed, pronounced tendencies seldom show the expected consistency in their judgments: the greatest adherent of softness may none-the-less prefer the stiffer of two taffetas, or the lover of roughness may prefer the smoother of two satins. It seems impossible to maintain a consistent attitude toward any one separate attribute because of our observation already cited, that no one attribute is a discretely given unit, but appears always in combination with others to make up a certain type of fabric. O's sometimes gave spontaneous expressions to their inconsistencies, qualifying their likes or dislikes by adding, *"for this type of fabric."*

"In this type of material, I don't seem to like them so soft"
"Lighter, and for that type of material should be light"
"Has a more satiny finish and though I disliked satin before, I like the finish in this kind of material."

We may summarize these observations by saying that there was no *specificity* of judgment with regard to single attributes. In the case of the preferential judgments, this is strikingly

comparable to what von Allesch found in the aesthetics of color (4).

The Influence of Cognitive Factors

In many cases it was obvious that the feeling tone was not directly given with the tactile contact but was significantly influenced by factors of experience.

The most frequent of these factors is the *identification* of the fabric, by which is meant, strictly speaking, the O's ability to name the fabric. Being able to name a fabric seemed to imply also that the O could place the fabric with regard to its appearance and with regard to a specific purpose which it fills in the broad field of garments.

The rôle of identification is clearly indicated in the following examples which show the change of the O's liking for a fabric where everything else remains constant and only the identification is added.

(O asked on second visit to "see" a certain fabric again)
"This may be gingham. If it is a gingham I can see it. I like it much better"

(O was asked: "Would you like this for yourself?")
"Well, I'm trying to decide. If it's wool, I would; if it's silk, I wouldn't"

In our *preliminary procedure,* more than half the O's gave a different rating upon seeing the linen. Of these, sixteen referred to no attribute outside the tactile field beyond the identification of the fabric; seven raised their ratings simply because they now recognized it to be linen, and six raised their ratings because they saw some specific use for the fabric.

Very closely linked with the factor of identification, and according to our assumption, part of it, is that of *purposefulness.* Many individuals were incapable of making a purely aesthetic judgment and refused to compare fabrics as such without knowing the purpose for which they were intended. This attitude was fairly consistently shown in several cases and seemed to be a function of a highly practical type of per-

sonality. However, even the more aesthetically inclined O's were occasionally blocked in giving a hedonic evaluation when they could not place the fabric in this respect. To begin with, many O's replied to our first question "How do you like it?"—"It all depends what for;" or "For what purpose?"

> "Don't like it because I can't think of any use for it"
> (O rejected a fabric) " . . . chiefly because I just can't place it and that gives me an awfully unpleasant feeling"

Another factor which we assume to be included in identification is that of *appearance*. The way a fabric feels and the way it looks seem to be inextricably connected. Even the comments of our four supplementary blind subjects bear this out. They speak continually of "seeing" the fabric:

> "May I see them together?"
> "Some people like black satin but I don't"
> "I always think of them in terms of durability and the way they would look"

Other O's are quoted:

> "I get the idea it has a glossy finish"
> "Women see through their fingers. I'm afraid I rank things according to colors I have in my mind's eye"
> "When I feel a thing, I see it walking down the street"
> "Distinctly gives the feeling that it's shiny . . . "
> "When I think how it would look, I'm not pleased with it because it is just a plain weave"

Although we asked our O's for no introspective reports, we have sufficient evidence in our data to suggest the underlying processes for judgments such as those mentioned above. They may possibly be best described by borrowing the phrase "categorical imperative," and are analyzed as follows. Once a fabric is perceived as such, it is classed in a category either in the sense of a tactile-kinaesthetic configuration or in the further sense of "identification." This type then, in turn, makes certain demands upon the individual attributes which go to make it up,

such that the O judges the attributes according to the way in which they measure up to the standards of the type. In other words, the O's experience a compulsion arising from recognition of the whole which affects their judgments of the parts—hence the phrase "categorical imperative." The following examples are to the point:

> "More body, heavier, rougher—the way it should be in character"
> "More what taffeta should be in flexibility and lightness"
> "Rough to the touch. (Do you like that?) No, not in a silk"
> "This would be a heavy chiffon but a light crepe."

In this type of comment, of which we had a great number, two factors were constantly stressed: one, the relativity of judgments of individual attributes as the fabrics differ; and two, the hortatory character of judgments according to which things are or are not as they *should be*.

An interesting question which arises in this connection, and to which we have no answer, is whether this tendency originates in actual compulsions of the sensory field or in experience; further, what may the relative importance of these two sources be?

Aesthetic Considerations

Though the nature of our results precludes the possibility of establishing a hedonic hierarchy of attributes, they point to general hedonic principles which apply to the grouping of attributes. Two quite different principles for combining elements prove of aesthetic value. They are best described by a comparison with the way in which colors may be combined so as to bring about pleasant effects. The first of these is by combining *blending* tones and achieving a harmony; the second is by *contrasting* tones and achieving a balance.

In order to define blending and contrasting tones in the field of fabric touch, we shall use our original classification of relaxing and demanding attributes. Expressing this classification

in aesthetic rather than sensory terms, we should restate it as follows: Under *relaxing* are grouped all those characteristics which make a more direct aesthetic appeal and which immediately and naïvely evoke a pleasant response, such as smoothness, softness, lightness, sheerness, fineness, and delicacy in general. Under *demanding* are grouped all those characteristics which, in all probability, depend upon a certain sophistication, a basis of experience to evoke a pleasant response and are, therefore, not directly judged by aesthetic criteria, such as roughness, resistance to crushing, heaviness, thickness, coarseness, and genuine sturdy qualities in general. We consider blending attributes those which belong to the same category, with reference to relaxing and demanding, and contrasting attributes those which belong to different categories.

According to the two principles, pleasant fabric effects can be achieved firstly by combining all relaxing elements or all demanding elements; secondly by combining certain demanding with certain relaxing elements in a balanced relationship. Further, it would seem that, with all fabrics rated along one axis having the most relaxing at one extreme and the most demanding at the other, the blending principle is most effective when applied to fabrics toward either end, and the contrasting principle when applied to fabrics in the middle section of the scale. In other words, where a fabric is endowed with relaxing attributes to such an extent that the majority of individuals feeling it, place the fabric as a whole in the relaxing category, the more it possesses of these characteristics the better. (Compare Table IV and discussion above.) A chiffon is liked not only because it is smooth and soft and light, but also because it is limp, fragile, bodyless, will fly away in the wind, etc.—the very characteristics which would make another fabric *dis*liked. On the other hand, a heavy, thick, rough, coarse fabric is liked, when it is liked at all, simply because these attributes are all in keeping with one another.

> "A thing as thin as this shouldn't weigh so much . . . should float more"

"In a light-weight fabric, I prefer fine weaves, finer threads, finely woven together"

"Rougher and more definite. If it *is* rough, I like it real rough"

"A cotton, I think. I like it. I like rough cottons and wools. It's heavy . . . it's roughness would look nice . . . the weave itself is coarser"

Effective as the principle of harmony proves, too much blending at either extreme is likely to be unpleasant, although not equally unpleasant for the relaxing and demanding ends. Most everyone would reject a fabric which synthesized all demanding attributes to a marked degree. Such a fabric would approximate burlap. If, in accordance with the principles of contrast, one factor is changed whereby softness is substituted for stiffness, the synthesized fabric would resemble a scratchy tweed and would probably find some admirers.

"Kind of scratchy, but all right for tweeds"

Should we go a step further and subtract the scratchy element leaving the surface still rough, the weave coarse, and the fabric heavy, we should arrive at something like a homespun and that would undoubtedly have a pleasant appeal for many O's.

Fewer O's would object to a synthetic fabric combining all relaxing attributes although, theoretically, such a tactile experience would correspond to a gastronomic experience—too sweet or too rich.

The following selection of comments further shows the tendency for balancing the demanding with the relaxing elements along the middle range of the scale.

"Not much body but it doesn't matter for this type of material because it's firm enough not to be flimsy"

"I think I like it because it feels both soft and firm at the same time"

"Rather soft and still has that springiness. Sort of rough to the surface but combined with softness which makes it more pleasant"

"The surface is smooth but still has irregularities that make it interesting to feel and gives you a contrast when you feel it"

"Sort of silky roughness which I think is a nice feeling"
"Nice dainty rough finish"

Again in connection with these aesthetic considerations the question arises as to the relative importance of experience. Is appreciation for the demanding end of the scale entirely or in great part learned? Would extremely naïve O's show preference for fabrics "blended" out of all relaxing attributes?[3] Would children or men show a difference in preferences from those shown by our women O's? The results of further experimentation are needed before these questions can be answered.[4]

IV. RELIABILITY

Chance Factors Influencing Judgments

We must assume that every choice is the resultant of two forces: a constant factor of the O's taste, and chance factors in the present situation. We can estimate the relative importance of these factors by the use of certain controls which consist in repeating choices to see what part of them remains constant. We devised several checks which we used occasionally and with a part of our O's.

The most striking piece of evidence for the constancy of the group judgment is the similarity of the order of preference for the fabrics which were used both in the comparisons of pairs and one of our subsequent series. In the first, the order is obtained by the average ratings given the fabrics. In the second, where the O was given 7 fabrics to choose between for first, second, third, and last choices, the order is obtained by the average choice assigned to each fabric. In other words, the first order is one of rating, in which individual items are rated in accordance with a fixed scale; the second order is one of

[3] We found no relation between relaxing or demanding preferences and the economic, educational, or age status of our O's.

[4] A comparison is suggested in the field of music with Moore's genetic theory of consonance, according to which the appreciation of dissonance increases with sophistication (5).

TABLE VI

Comparison of Preferences for Fabrics in Series I and II

FABRICS	SERIES I AVERAGE RATING 1–5	SERIES II AVERAGE RANKING 1–7
Velvet	4.66	2.45
Flat	4.32	3.67
Satin	3.96	4.00
Sheer	3.28	4.87
Taffeta	3.87	5.17

ranking, in which individual items are ranked in comparison with one another. These figures are presented in Table VI and show the identical order of fabrics for the fabrics in these two differing set-ups, spaced approximately forty-five minutes apart in point of time.

A similar relationship between rating and comparative choices is found in the results of the comparisons alone. In general, the higher the average rating for a fabric by all the O's, the greater was the percentage of choice for the fabric. In Table VII, the order for the ratings is not based on the actual

TABLE VII

Comparison of Rayon Ratings and Choices

FABRIC	RATING ORDER	PERCENTAGE OF CHOICE ORDER
Medium crepe I	1	2
Light crepe	2	1
Satin	3	4.5
Flat print	4	3
Heavy sheer	5	6
Medium crepe VII	6	4.5
Heavy crepe	7	9
Thin sheer	8	8
Taffeta	9	7
Velvet	10	10

ratings but on the difference in ratings between the two members of the pair, beginning with the greatest differences in favor of the rayon and ending with the greatest difference in its disfavor.

For 31 O's the first pair of fabrics was presented again at the end of the series with the same instructions as for any

other pair. This repetition took place approximately after a lapse of half an hour. The ratings given the fabrics on the first and on the repeated presentation, shown in Table VIII,

TABLE VIII

Ratings for Pair I on First and Repeated Presentation

	RAYON	SILK
First presentation	3.7	2.98
Repeated presentation	3.75	3.08

are practically identical and demonstrated the constancy of the group judgment.

The difference in choices made on the first and repeated presentations of the same pair is also an indication of reliability. Of the 31 O's, 24 chose rayon on the first, and 23 chose rayon on the repeated presentation.

In the cases of 50 O's a pair of *identical fabrics, i.e.,* two pieces of the same cloth, was presented with the same procedure accorded the other pairs in Series I. It was inserted between pairs VIII and IX. The fabric was a wool serge—purposely different in fibre and character from the fabrics properly included in Series I. The subjects made the following judgments:

Preference for Piece 1	13	26%
Preference for Piece 2	10	20%
No Preference	27	54%

The average number of "no preference" judgments for all the other pairs is 16.3 or 16.6 per cent with a range from 2 per cent to 25 per cent. It is, therefore, readily apparent that the identical pair yielded a vastly greater number of "no preference."

With reference to chance factors which may alter judgments from one repetition to another, we should like to point to the time factor as being of psychological interest. Apparently it has an influence on the data, lessening the number of comments made and increasing the number of indecisions, *i.e.,* cases of no preferences the later a pair is presented. There is a

rank correlation of .55 between the amount of comments made on a pair of fabrics and the presentation rank of that pair in the series. Perhaps the effect of fatigue (or boredom or satiation) is better shown by the increase in the number of indecisions with time. There is a rank correlation of .77 between the number of indecisions and the presentation rank of the pair in the series. Of course, this figure is necessarily connected with that just given for, obviously, the more indecisions, the less reasons for choices.

Judgments as Expressions of Personal Tendencies

The question can be asked to what extent the choices are controlled by the nature of the fabric and to what extent subjects have definite preference, as *e.g.*, for light fabrics or for rayon, offered for choice. The answer to this question lies in the following analysis:

One thousand choices were made by the 100 O's on the ten pairs. If 160 of those choices were, for instance, motivated by a certain comment then the probability of such a comment is .16 under the conditions of this experiment. With the help of the binomial formula, it can be computed what chance distribution of these comments among the O's should be expected. The following table shows in its first line that for a probability .16, about seventeen subjects should make no comment of this type, thirty-three should make one such comment, twenty-seven should make two, and so on.

TABLE IX

Number of L Comments, p = .16

	0	1	2	3	4	5	6	7	8	TOTAL
Theoretical Dis.	17	33	29	15	5	1	0	0	0	100
Actual Dis.	33	27	18	8	6	4	2	1	1	100

The second line of the table gives the actual distribution of L comments in our experiment. Thirty O's chose the fabric because it is light, twenty-five chose one, sixteen chose two, and so on. The table shows that the number of people who

made no L comment or who made four or more comments is much greater than should be expected by chance.

In compensation of that, the number of O's making an average amount of comments is smaller than chance would let us expect. It is, therefore, justifiable to say that there exists a definite personal tendency for choosing or not choosing a fabric because it is light.[5] The subjects are inclined to make either no L comments at all, or to make them rather frequently.

In the same way it can be shown that the selection of fabrics on the ground of their being heavy or having demanding qualities (smooth, soft, etc.) is partly due to "personal tendency." By that is meant here again that the distribution of the number of such comments per O is beyond any reasonable chance different from the theoretical distribution which one might expect on the ground of the average probability of such choices.[6]

As a further example, the actual distribution of X comments is compared with the chance distribution for a probable ♯6, 600 choices out of a thousand having been motivated by a preference for relaxing attributes.

TABLE X

Number of X Comments, p = .60

	0	1	2	3	4	5	6	7	8	9	10	TOTAL
Theoretical Dis...	0	0	1	4	11	20	25	22	12	4	1	100
Actual Dis.	0	2	3	9	17	11	11	18	14	12	3	100

Again the number of people making less than four or more than eight comments is much greater than should be expected by chance and has to be accounted for by an assumed tendency to prefer or not to prefer X qualities different for different groups of people.

It was especially interesting to test if preferences for rayon are, themselves, a personal tendency in the statistical sense used so far,—in other words if there is something like a rayony quality of which the O would either be particularly

[5] The χ^2 test for goodness of fit shows that our actual distribution of comments deviates from the theoretical distribution far beyond any reasonable chance probability.

fond or the opposite. The probability of a rayon choice is, as we have seen, .54. For such a probability, the distribution of choices to be expected is given in the first line of the following table. The second line gives the actual distribution of the rayon choices. It is easy to see how little the two distributions differ. Therefore, there is no "rayony" quality in the sense that there is something like a demanding and relaxing quality in a fabric.[6]

TABLE XI

Number of Rayon Choices, p = .54

	0	1	2	3	4	5	6	7	8	9	10
Theoretical Distribution	*0*	*0*	*3*	*8*	*17*	*24*	*23*	*16*	*7*	*2*	*0*
Rayon Choices	0	1	5	11	12	22	25	12	8	4	0

In spite of the fact that we cannot find, statistically, any personal tendency to perceive and prefer rayony quality, the subjects have a very definite rational attitude. Although we found that markedly more rayon fabrics were selected than silk fabrics, 61 per cent of the subjects declared upon questioning that they disliked rayon and only 14 per cent liked it. A further test showed that the correlation between attitude and choice is exactly zero.

SUMMARY

1. A series of paired comparisons, consisting of 10 pairs with a silk and a rayon member each, was presented to 100 blindfolded women O's.

[6] For readers not trained in this type of analysis, it should be pointed out that the question of an especially great or an especially small tendency of an individual to select rayon fabrics would be independent of the average amount of rayon choices in the whole group. We found a definite preference for rayon in the whole group but not personal tendency beyond that in single individuals. It could as well have been that a group preference for silk goes together with the personal tendency toward "rayony" choices. Any combination of a personal tendency and a group preference is possible. For the theory of Lexis underlying this analysis see the book of Rietz (6).

2. It was found that the rayon member was preferred in 64 of all the choices.

3. Sixty-four per cent of the O's gave the same preferences on repetition after half an hour; the judgment of the group remained practically constant.

4. For purposes of systematization it appeared best to classify tactile attributes according to two qualitative dimensions, *surface* and *depth*.

5. All our theoretical conclusions must be stated in terms of combinations of attributes rather than single attributes, inasmuch as the latter were rarely experienced and reported as such.

6. There is no specificity of judgment, either tactile or hedonic, with respect to any one attribute; all judgments are influenced by the configuration which goes to make up the fabric and any cognitive factor which may be connected with it through experience.

7. There is clear indication for two types of pleasantness in tactile experiences which we have termed the *relaxing* and the *demanding*.

8. Pleasantness in a fabric may be achieved either by *blending* or by *contrasting* attributes.

9. The reliability of the judgments for the whole group of O's tested repeatedly is high.

10. Applying a statistical test upon the distribution of comments they turn out not to follow a chance distribution. One has to assume, therefore, that the probability of choices is different for different groups of O's. There seems to be personal tendencies in the field of tactile preferences but nothing can be said so far on their nature or origin.

REFERENCES

1. Beebe-Center, J. G. The Psychology of Pleasantness and Unpleasantness. N. Y. D. Van Nostrand Co., Inc. 1932.
2. Katz, David. Der Aufbau der Tastwelt, Z. C. Ps. Ergaenzungsbd. 11. 1925.
3. Major, D. R. On the Affective Tone of Simple Sense Impressions. Amer. J. Psychol. 7, 1895, 72.

4. VON ALLESCH, G. J. Die Aesthetische Erscheinungsweise der Farben. Psych. Forschung, VI, 1925.
5. MOORE, HENRY THOMAS. The Genetic Aspect of Consonance and Dissonance. Psychol. Monog., 17, #73, 16–68.
6. RIETZ, HENRY LEWIS. Mathematical Statistics. Open Court Publishing Co. 1936.

COMMENT BY MR. ALEXIS SOMMARIPA, DU PONT REPRESENTATIVE, UNDER WHOSE DIRECTION THIS INVESTIGATION WAS MADE.

It is well known that the making of textiles occupies about 25 percent of all people engaged in manufacturing. The value of textile products, even in 1933, was about $5,000,000,000.

There is no question that among the buying motives for textiles, the pleasantness to the touch plays a very important role. No man or woman, probably, would buy a garment without first putting a hand on it. Yet, notwithstanding the magnitude of the textile business or the importance of touch, there is practically no concrete information on the subject. It is true that in designing fabrics or garments, a great deal of attention is paid by the trade people to the "hand" of the cloths. But there are no clear-cut criteria established.

In the production of synthetic yarns and fabrics, the manufacurers can modify their products with regard to the resultant touch. It was logical, therefore, for rayon people to try to get some knowledge as to the preference of the consumers for fabrics having different types of "hands." The study carried out by Doctors Lazarsfeld and Ripin for the Du Pont Company has been a pioneering one. It is of interest that all through the study, they have reported a tremendous liking on the part of the subjects toward a type of rayon fabric which was made for the first time and a dislike for another competitive fabric. Since then, the judgment of the few people was confirmed by the country at large. The 10-yard sample has grown into a consumption of 115,000,000 yards during 1936, displacing the disliked competitive product.

The study demonstrated the possibilities of investigations of consumer reactions to the sense of touch. It has shown, how-

ever, that like any other scientific investigation, it is essential to make comparisons between the items which have preferably only one variable. On the other hand, it has also been shown that the preference for the touch of a fabric depends upon other composite factors, such as its use, weight, and appearance. The study seemed to have indicated that, at a given period, the reactions of a group of consumers to certain criteria are quite similar. I think, however, that in the course of many years, some of the sensations that are considered agreeable will be considered disagreeable, depending upon the general living conditions and mental attitudes.

THE FACTOR OF MECHANICAL ARRANGEMENT AND TYPOGRAPHY IN QUESTIONNAIRES

HERBERT A. TOOPS
Ohio State University

THE common tendency in the construction of questionnaires is to weight heavily the demands of good rhetoric and good printers' conventions, add to that an abbreviated and stilted type of English, and call the result a good job.

In the construction of many questionnaires and blank forms we have come to the general conclusion that a questionnaire most hopefully and most helpfully is to be viewed in its light of being a stimulus for securing a certain end result, thus necessitating the presupposition that everything should be rigidly subordinated to that end. Accordingly, we no longer hesitate to use redundancy, for example, if only that will promise to make our questions more clear to the respondent; we likewise conclude that the rules of rhetoric are important in questionnaire construction only insofar as their use or their appropriate disregard leads to a better understanding of our intent on the part of the recipient; and, when printing a questionnaire we fight an eternal warfare with the printer and particularly the compositor, thus earning no doubt, by our disregard of his most sacred conventions, his haughty contempt. In addition to all of these sins recently in the construction of a personnel form, the answers to which were to be read by three departmental members for upwards of some one hundred candidates annually, we acquired a new "vice" which promises to be so useful that we here bring it to the attention of others who may be interested.

The matter best may be illustrated, perhaps, by taking the reader through the "thinking steps" in the evolution of the new technique. We were in the process of revising a departmental personnel form, an application blank which all applicants for assistantships, fellowships, scholarships and instructorships in the Department of Psychology at the Ohio State University annually fill out as a preliminary to being considered as applicants.* In this form, Section 8 dealt with "Organizing and Executive Ability." The original questions, as printed in the first edition, are shown in Fig. 1.

FIGURE 1

The Original Four Questions and Their Arrangement

1. What projects, clubs, and societies have you *organized?*

.....

2. What projects, clubs, and societies have you *managed?*

.....

3. What positions involving executive responsibility have you held?

.....

4. Give any other evidence of your executive ability (e.g., executive experience in business).

.....

The questions being somewhat inadequate to call forth all of the possible responses of the applicant in regard to his organizing and executive ability, it was decided to add additional questions, with the result that in conference the following questions were propounded:—

1. What is the most important responsibility you have ever held?
2. How many people have you supervised simultaneously and successfully?
3. Have you ever settled disputes; built up morale?

* Toops, Herbert A.: "The Selection of Graduate Assistants." The Personnel Journal, Vol. 6, No. 6, April, 1928, pp. 457–472.

The three new questions were added at the foot of the previous list after moving the fourth of the original questions (See Fig. 1) to the foot of the list as the concluding question. That is to say, the three new questions were given numbers 4, 5 and 6, and original question No. 4 was made No. 7 in the list.

In the ensuing discussion the following points, in serious criticism of the original arrangement of Fig. 1, were brought out:—

1. In scanning the answers of applicants to the questions of Fig. 1, it was rather difficult to tell to which question a given answer referred without first reading the original question. This becomes very laborious and boresome when one has upwards of a hundred such to read. Even if one reads the question—thus holding down markedly one's speed in inspecting the answers—eventually the questions became somewhat ineffective as a stimulus: one becomes hypnotized to them, so to speak and in so doing, becomes hypnotized as well to the sense of the answers, perhaps.
2. The arrangement of the questions does little to facilitate the reader's legitimate desire to identify the stimulus question without the necessity of reading the whole "question."
3. The answers take a great deal of eye-movement in order to comprehend the whole.

Accordingly, it was decided to capitalize the important key-words, no attempt being made, at this time, to solve Number 3 of the above criticisms. The resulting typewritten copy is given in Fig. 2.

It was clear, from an inspection of the typed copy (Fig. 2) that the new blank, when printed, would be quite as bad, if not worse, than the old in respect to criticism No. 3. Careful inspection of Fig. 2 showed that by shifting Question 4 to first-place, Questions 3, 1, and 2, to second, third and fourth places, respectively, and then dividing each question at the points indicated by the paragraph signs in Fig. 2, it would be possible

FIGURE 2

The second Step in the Evolution of the New Arrangement after three New Questions were added

1. What projects, clubs and ¶SOCIETIES have you ORGANIZED¶

2. What projects, clubs and ¶SOCIETIES have you MANAGED¶

3. What positions involving ¶EXECUTIVE RESPONSIBILITY have you held¶

4. What is the MOST IMPORTANT RESPONSIBILITY ¶you have ever held¶

5. How many PEOPLE have you SUPERVISED ¶simultaneously and successfully¶

6. Have you ever SETTLED DISPUTES; ¶built up MORALE¶

7. Give here any OTHER EVIDENCE of your ¶executive ability (e.g., executive experience in business).

to shift the significant key words into a mechanical arrangement such that the reader by a very simple eye movement would easily become cognizant of the *meaning* of a given question by reading only the capitalized words in a comparatively simple eye-movement. The printed end result is Fig. 3.

The consensus of judgment of the three readers is that the new form is immensely superior to the old. Since it is they who ought to be pleased, certainly no elaborate statistics need be adduced to prove the worthwhileness of the new technique.

As a second example of the same technique, we present, in Fig. 4, four questions, similarly revised, appearing in a section of the same revised questionnaire, dealing with the matter of the mechanical ability of the applicant.

FIGURE 3

The Revised Section, as Finally Printed

Organising and Executive Ability

What is the MOST IMPORTANT RESPONSIBILITY you have ever held?	
What POSITIONS involving EXECUTIVE RESPONSIBILITY have you held?	
What projects, programs, clubs, and SOCIETIES have you ORGANIZED?	
What projects, programs, clubs, and SOCIETIES have you MANAGED?	
How many PEOPLE have you SUPERVISED simultaneously, and successfully?	
Have you ever SETTLED DISPUTES; built up MORALE?	
Give any OTHER EVIDENCE of your executive ability? (e.g., executive ability in school or business)	

FIGURE 4

A Similar Example from the Section Dealing with Mechanical Ability

What ENGINEERING EXPERIENCE if any, have you had?	
What devices or APPARATUS have you DESIGNED?	
What devices or APPARATUS have you CONSTRUCTED?	
Amount of MACHINE SHOP and related EXPERIENCE?	

VALIDATING TWO TESTS FOR CENSUS ENUMERATION

HERMAN A. COPELAND
Cincinnati Employment Center

APPROXIMATELY 750 applicants for various jobs on a Regional Census conducted by F.E.R.A. were tested by means of two clerical tests, the Otis Self-Administering Test of Mental Ability Higher Form D with a 30-minute time limit and the Minnesota Vocational Test for Clerical Workers. These tests are now being used in the Examining Division of the Cincinnati Employment Center as part of the routine registration procedure for clerical applicants. Nine months after the census had been in operation, efficiency ratings on many of these registrants were secured and were compared with their test scores. The validity of these tests has been determined for the following groups: supervisors, clerks, and enumerators.

I. Thirty-eight *supervisors* were selected by the director of the project from the entire group of applicants. Their average test scores on the two tests and the correlations between the scores on these two tests and two distinct efficiency ratings are shown in the first table. The correlation of .58 between the first and second rating is typical of the agreement between separate ratings made by an inexperienced rater. This reliability lowers the correlation between test scores and efficiency on the job.

TABLE I
Validity of Tests for Selecting Supervisors

	AV. SCORE	CORRELATION WITH FIRST RATING	CORRELATION WITH SECOND RATING
Otis S–A Form D	41	.22	.23
Minnesota Clerical	199	.12	.28

II. There were fifty-six *clerks*. The average test scores and the correlations between the test scores and their efficiency on the job are shown in the second table. The clerical test seems superior to the Mental Ability Test with regard to clerks.

TABLE II
Validity of Tests for Clerks

	AV. SCORE	RATING
Otis S–A Form D	33	.07
Minnesota Clerical	198	.18

III. Two studies were made of the *enumerators* who went from house to house collecting the basic census data. The first study of 145 enumerators made use of supervisors' efficiency ratings, and the second study of 276 enumerators made use of daily production records (number of contacts). Table III shows the average scores and correlations between scores and ratings or production records.

TABLE III
Validity of Tests for Enumerators

	AV. SCORE	1. CORRELATION WITH RATING
Otis S–A Form D	34	.14
Minnesota Clerical	181	.20
		2. CORRELATION WITH PRODUCTION
Otis S–A Form D	35	.18
Minnesota Clerical	176	.01

These workers had been chosen from the best applicants available for F.E.R.A. work and the choice was based upon education, experience, test scores (including another test designed especially for this specific problem), and interviewers' impressions. From the results of previous investigations it was believed that those applicants receiving low scores on all tests would be unsatisfactory. Accordingly, they were usually

rejected; as a result the validity is lowered because of the resulting restricted range of ability.

These data may be evaluated in the following manner: most supervisors scored higher on both tests than clerks and enumerators; although some supervisors received lower scores than some clerks or enumerators. On the clerical test the clerks averaged higher than the enumerators, and only very slightly lower than the supervisors. From this it can be inferred that the high clerical test scores had been helpful in selecting clerks.

These correlations based exclusively upon test scores, which are only *one* part of the registration record, are not as high as have been reported in many textbooks and journals. In light of the complexity of human conduct, some of the growing reports in textbooks are so unusual that one cannot hope to duplicate them in daily practice. In addition the practical value of low correlations such as 0.10, .20, or .30 has been minimized in testing even though agricultural statisticians well recognize their usefulness.

Although these test scores are one and only one part of the total record, they do add to it information of such value that the scores alone would select *some* good employees; and along with other information they aid in selecting more good employees and eliminating more poor employees.

DO YOU KNOW THE MULTIPLICATION TABLE?

RICHARD J. TRIPLETT
University of Denver

WHAT is the Multiplication Table? How much of it is one supposed to know? When does one know it, or a portion of it? Do students from school of large cities know it better than students from schools of the smaller cities or rural districts? Do boys know it better than girls? If a business man were to demand perfection—one hundred per cent accuracy—what per cent of our high school graduates could he hire? If he were satisfied with 98 per cent, 95 per cent, or 90 per cent accuracy, how many would be eligible? How about the "slow and accurate" theory, or fallacy?

Data from four hundred nineteen commercial school students (high school graduate) were collected in an effort to cast light on some of these questions. "Just answers," determined by a preliminary survey conducted to make this study possible were given to others.

"JUST ANSWERS"

What is the Multiplication Table? "It is up to twelve times twelve, according to the arithmetic I studied." That is the answer received from many teachers, business men, and other persons. In nearly all cases persons giving this glib reply had forgotten the name of "their arithmetic." Others said, "It ought to be up to twenty-five times twenty-five." Such persons invariably admitted that they did not know it up to that point.

How much of it is one supposed to know? "Up to twelve times twelve"—that may not be a logical or scientific answer but it is surely the traditional and accepted standard.

When does one know it, or a portion of it? In constructing a test to answer certain questions, it was discovered that ordinarily we do not know how much we know or when we know items in the table. Twelve times twelve was found not to be the limit, even though it is seldom studied beyond this point. By a process of unconscious learning, some of the table up to twenty-five times twenty-five, probably still higher, appears to be known to almost everybody. For example, high school graduates were found to be as likely to know the product of three times fifteen as they were to know the product of three times twelve. When asked such items as three times seventeen the answers were often written as quickly and as accurately as answers to items like three times eight. It was noted, however, that the figures in the answer were sometimes written from right to left instead of left to right. Although the answer was written without apparent hesitation the subject probably did not know the answer but "multiplied it out" so readily that he did not realize what was happening. Consequently, it is doubtful if one can tell "how much he knows" or "when he knows."

THE TEST

After making a preliminary survey it was decided to construct a test containing all the possible combinations up to twenty-five times twenty-five. In mimeographing the test a few combinations were omitted, so it actually contained 611 combinations. The combinations were presented in random order as indicated by a section of the test presented below.

Instructions: Fill in the product of the following numbers as rapidly as possible. Skip if you do not know.

12×11	4×16	9×18	5×19 .....
6×7	14×9	8×15	7×11 .....
3×7	6×13	12×13	3×14 .....
8×6	9×15	21×23	3×18 .....
10×13	12×15	9×6	9×8 .....

A group of twenty-five cooperative students and teachers of varying ability was selected and the test was tried out so as to

determine a time limit. They were instructed not to multiply, but to skip items that they did not know. However, the tendency to multiply is so strong that even this group probably failed to follow instructions to the letter—one can't be sure whether he knew or whether he multiplied in certain instances.

The time required by this group was from fifteen to twenty-one minutes. In view of this finding, the time limit for the test was set at twelve minutes. It was felt that students attempting to increase their scores by multiplying would automatically be penalized by this limit.

When the test was given the students were given both written and oral instructions to skip items that they did not know, not to multiply, and that their score would automatically be lowered if they attempted to multiply. Probably every one violated these instructions. It is perhaps impossible to do otherwise. However, such violations would not seem to invalidate the test—it is still a multiplication test, and it is believed to be an indicator of one's knowledge of the multiplication table.

THE GROUP TESTED

Four hundred nineteen students of the Barnes School of Commerce of Denver, high school graduates from many high schools in the Rocky Mountain region, were tested. The intelligence of the students was not determined, but this school is known to attract a high type of student. Six hundred students from the same school were tested and found equal to Liberal Arts and technical college students.[1]

For the purpose of this study the students were divided into four groups: Group A, students who received their education in the schools of cities having a population of 50,000 or over; Group B, students who received their education in towns having a population of 2,500 to 50,000; Group C, students who received their education in communities having a population of under 2,500; and Group D, students who attended schools

[1] Intelligence of Commercial College Students, by Richard J. Triplett, The Personal Journal, Vol. 13, No. 2, August 1934.

in two or more of the above classifications, a mixed group. Table I indicates the members in each division.

TABLE 1

Distribution of the four hundred nineteen students tested for knowledge of the multiplication table

POP. OF COMMUNITY WHERE THEY ATTENDED SCHOOL	NUMBER OF STUDENTS		
	Male	Female	Total
Group A Over 50,000	27	50	77
Group B 2,500–50,000	12	52	64
Group C Under 2,500	32	77	109
Group D Mixed*	55	114	169
	126	293	419

* This group received training in two or more of the indicated population areas.

RESULTS OF THE TEST

The test was scored in two ways: First, the number of items correct; and, second, the per cent of attempted items correct. Hereafter, the results will be spoken of as "Number correct" scores, and "Per cent correct" scores.

The "Number correct" scores ranged from 61 to 403. The mean for the group was 215.07 ± 1.87. The median was 214.10. The Standard Deviation was 60.11.

Table 2 gives a summary of the findings.

TABLE 2

Summary of findings from "Numbers Correct" scores on Multiplication test

NUMBER	MEAN	MEDIAN	S.D.	LOW SCORE	HIGH SCORE	Q_1	Q_3
419 Students	215.1 ± 1.9	214.1	60.1	61	403	177.2	256.0
293 Girls	214.6 ± 2.3	214.3	58.1	61	372	175.7	255.0
126 Boys	217.9 ± 3.8	213.5	63.2	70	403	179.8	258.6

In an effort to determine the effect of different school advantages, these data were broken down into four groups as indicated in Table I. Table 3 shows a summary of results for

each division. Some of the groups are small, and the corresponding results are therefore questionable, but it is clear that a line may be drawn at about the 2,500 population mark.

Using medians as criteria, girls in communities having a population of 2,500 or over seem to be poorly trained in regard to the multiplication table when compared with the boys of similar communities; in communities of under 2,500 population, the girls appear to be receiving superior training. The boys and girls that received training in two or more of the population areas in Group D exhibited about equal ability.

TABLE 3

Group results for "Number Correct" scores on Multiplication test

	MEAN	MEDIAN	Q_1	Q_3
Group A (over 50,000 pop.)				
27 Boys	218.1	216.1	181.3	257.8
50 Girls	195.0	193.3	165.6	223.6
Group B (2,500 to 50,000 pop.)				
12 Boys	247.9	237.5	208.3	275.0
52 Girls	213.0	210.7	162.5	258.3
Group C (under 2,500 pop.)				
32 Boys	203.1	192.9	150.0	241.7
77 Girls	235.6	236.5	200.4	271.3
Group D (mixed)				
55 Boys	219.8	216.3	183.5	260.2
114 Girls	209.7	213.0	172.1	246.7

The "per cent correct" scores ranged from 52 to 100. Four girls received a perfect score. The mean for the group was $95.0 \pm .15$. The median was 96.0. The standard deviation was 4.9. In Table 4 the findings are summarized.

The "per cent correct" data were broken down into population groups and is shown in Table 5. Whatever may cause the apparent sex differences found in the different training groups based on "number correct" scores, it did not affect "per cent correct" scores similarly. Using either means or medians as criteria, the boys and girls seem to be equal in the

TABLE 4

Summary of findings from "per cent correct" scores on the Multiplication test

	MEAN	MEDIAN	S.D.	LOW SCORE	HIGH SCORE	Q_1	Q_3
419 Students	95.0 ± .15	96.0	4.9	52	100	93.2	97.9
293 Girls	94.9 ± .20	95.9	5.1	52	100	93.1	97.0
126 Boys	95.3 ± .25	96.2	4.2	72	99	93.4	98.0

matter of accuracy, except in the large city groups. In this group the figures are slightly higher for the boys, but the number of boys in the group is too small to make the differences meaningful.

TABLE 5

Group findings from "per cent correct" scores on Multiplication test

	A. M.	MEDIAN	Q_1	Q_3
Group A (over 50,000 pop.)				
27 Boys	95.1	96.9	93.7	99.2
50 Girls	92.9	94.5	92.3	96.7
Group B (2,500 to 50,000 pop.)				
12 Boys	95.5	96.2	92.0	98.0
52 Girls	95.2	96.3	93.5	98.6
Group C (under 2,500 pop.)				
32 Boys	95.4	96.0	93.2	97.6
77 Girls	95.8	96.4	94.6	98.0
Group D (mixed)				
55 Boys	95.4	96.1	93.6	97.8
114 Girls	95.1	95.8	92.7	98.2

If a business man were to demand perfection—one hundred per cent accuracy in the matter of multiplication—what per cent of our high school graduates could he hire? If the four hundred nineteen examined are truly representative, one would conclude that only one per cent would be eligible. If a 98 per cent standard were acceptable, 24 per cent would be eligible; a 95 per cent standard would make 63 per cent eligible; and

by a 90 per cent standard 91 per cent would be eligible. Table 6 gives cumulative totals of these percentages.

TABLE 6

Cumulative percentages based on "per cent correct" scores for 126 boys and 293 girls, determined by the multiplication test

	TOTAL		MALE		FEMALE	
	Correct	Cum.	Correct	Cum.	Correct	Cum.
98–100	24.3	24.3	25.4	25.4	23.9	23.9
95– 97	38.2	62.5	40.5	65.9	37.2	61.1
92– 94	20.3	82.8	17.5	83.4	21.5	82.1
89– 91	9.8	92.6	10.3	93.7	9.6	92.2
86– 88	4.3	96.9	4.0	97.7	4.4	96.6
83– 85	1.2	98.1	.....	97.7	1.7	98.3
80– 82	0.7	98.8	1.6	99.3	0.3	98.6
77– 79	0.2	99.0	.....	99.3	0.3	98.9
74– 76	.....	99.0	.....	99.3	.....	98.9
71– 73	0.2	99.2	0.8	100.0	.....	98.9
68– 70	0.2	99.4	.....	.....	0.3	99.2
65– 67	.....	99.4	.....	.....	.....	99.2
62– 64	0.2	99.6	.....	.....	0.3	99.5
59– 61	.....	99.6	.....	.....	.....	99.5
56– 58	.....	99.6	.....	.....	.....	99.5
53– 55	.....	99.6	.....	.....	.....	99.5
50– 52	0.2	100.0	.....	.....	0.3	100.0

Is there a tendency for the "slow" students to be more accurate? This study, in accord with many other studies, shows that students are slow and inaccurate rather than slow and accurate. Instead of the coefficient of correlation between the "Number Correct" scores and the "Per cent correct" scores being negative, it is positive and fairly high—.47 ± .03.

SUMMARY

1. An analysis of the test papers shows that errors in items below twelve times twelve are nearly as frequent as in items above that point.

2. The population of the community in which a student receives his education appears directly or indirectly to have

an effect upon his capacity to multiply. The city boys are more rapid than city girls.

3. Accuracy in multiplication was not affected by the place where the student received his education.

4. The efficiency of our high school graduates is high or low according to the standard set. If one should be "letter perfect" on the multiplication table our students appear feeble indeed. If, under stress, not more than 95 per cent accuracy should be expected, about two-thirds of the students tested could qualify.

5. Students slow in the matter of multiplication are definitely not more accurate.

NEWS AND NOTES

The Eighth Annual Meeting of the Association of Consulting Psychologists will be held May 7–8, 1937, at New York State College for Teachers, Albany, N. Y. Friday evening, at 7: 30, a conference will be held in the State Education Building for the purpose of organizing a New York State Association of Psychologists in Clinical and Applied Fields under the direction of the New York State Committee: Charles A. Drake, Warren W. Coxe, and John G. Jenkins, Chairman. At 8: 00 P. M. an informal social meeting will be held in Hotel Ten Eyck.

On Saturday, May 8, the forenoon will be devoted to meetings in Educational, Industrial and Clinical Psychology. The afternoon program will consist of a symposium, panel discussion and round table conferences followed by the annual dinner at the Roger Williams Restaurant at 6: 30. Addresses will be given following the dinner by A. R. Brubacher, President, New York State College for Teachers, on "Some Observations of an Educational Administrator on the Functions of the Psychologist." This will be followed by an address by the retiring President of the Association, Gertrude Hildreth, on "The Psychologist's Part in Guidance and Therapy."

Other speakers on the various programs include Dr. Douglas Freyer, James Q. Holsopple, Emily L. Stogdill and Julius B. Maller.

Dr. Sidney L. Pressey, Professor of Educational Psychology in Ohio State University, will be among the visiting instructors in the University of Hawaii summer session beginning June 28. Attended by about 1200, the University of Hawaii summer session boasts students from twenty-seven states and six foreign countries, representing several racial groups.

The Sixty-first Annual Convention of the American Association on Mental Deficiency will be held in Atlantic City, N. J., from May 5–8, 1937, with headquarters at Hotel Haddon Hall. The three-day program is replete with a wide variety of interesting sessions from the call to order by President Benjamin O. Whitten, Superintendent of the State Training School, Clinton, S. C., to the last discussion on Saturday, May 8. Many outstanding persons in the fields represented will appear on the various programs arranged. The President's Dinner will be held at 7: 00 P. M. on Friday followed by the President's address and a muscial program. This Convention should be of interest to a large number of persons as the cause, the care and the prevention of mental deficiency is one of our economic, social and scientific problems.

Recently this Journal received from Dr. Richard S. Schultz of the Psychological Corporation of New York City a resumé of the address of Dr. L. A. Lincoln, President of the Metropolitan Life Insurance Company, delivered at the conclusion of a three-day managerial convention of that Company. Mr. Lincoln stressed the importance of applying sound scientific management principles to the selection and supervision of agents. "After all," he said, "management means the selection of men of good character and potential ability, training them in the proper methods of work, supervising them carefully in their daily activities, and inspiring them to do the best work of which they are capable. . . . You may be interested to know that for the last four years we have been making scientific studies of men who have come into our business. This study has been made for us by the Psychological Corporation. From this study we are beginning to learn many things about the type of man who is most likely to succeed in our business. The reason we are trying to find some scientific method for use in the selection of men, is not alone because it is good for the company: it will weed out men who have no capacity for the business and in that way will be doing them a service.

"We now have six districts where men who have come into the business are being tested. The managers who are giving these tests very willingly offered to cooperate with us, and we hope in a year or so to have something very definite to show as a result of these experiments.

"It is of course obvious that no matter how good your selection may be, unless you recognize and assume responsibility for continuous training, your good selection will not amount to very much."

The International Congress of Psychology will hold its eleventh annual meeting at the Sorbonne in Paris, from July 25–31, 1937, with Dr. M. H. Pieron, College of France, President, and M. I. Meyerson, Secretary-General. The international executive committee is composed of Dr. Ed. Claparéde, of the University of Geneva, President; H. S. Langfeld, Princeton University, Secretary; and Drs. Myers (London), Pieron (Paris), Ponzo (Rome), Stern (Durham, U. S. A.) and Rubin (Copenhagen). A full and very interesting program is being arranged for those who attend, in addition to which will be many attractive excursions, receptions and fetes, to be announced in detail at a later time. Those wishing definite information concerning the Congress should write to the General Secretary, Dr. Meyerson, Directeur adjoint a l'Ecole Pratique des Hautes Etudes, Paris, France.

Expansion of the University of Oregon clinic for atypical children, a nationally recognized summer session feature for the past several years, into a year-round program is announced by Mrs. Elizabeth Montgomery, assistant professor of education. The work of the clinic will continue to be restricted to children rated as normal in intelligence, but who have reading difficulties.

The clinic serves a dual purpose, it renders service to hundreds of children, and provides working laboratory for the teaching of the methods employed. During past summers hundreds of teachers from all over the country have come to Eugene for training in this work, and now, under the all-year plan, advanced students in the school of education may take advantage of this opportunity to gain first hand knowledge of these latest methods in this work.

The university summer clinic, a pioneer in this type of work, was established in 1929, when Dr. Grace Fernald, Los Angeles psychologist, was brought to the Oregon campus. The late Dr. B. W. DeBusk, professor of education, worked with Dr. Fernald when the work was first undertaken here, and for the past several years had charge of the clinic. Under his direction hundreds of teachers have learned methods of teaching nonreading children, while pupils have been brought here for help from all parts of the state and even from distant parts of the nation.

In a recent report to the motion picture industry, Will H. Hays, President of the Motion Picture Producers and Distributors of America, declared that the 1936–37 season "not only disclosed new peaks of artistry in the production of notable film entertainment but brought a greater number of truly fine pictures to the screen than the industry has ever before produced in a similar period." The past year has proved that better pictures are erasing the distinction between any specialized audience and the number of possible patrons which fine entertainment can command. Many of the best pictures of the season were original screen plays. Whereas in 1935 the total of original screen material approved in Hollywood under the Motion Picture Production Code was 47 per cent, in 1936 the percentage increased to 67.82. Indicating the upward progress of the screen a number of musical biographies of the lives of great musicians, composers and artists are planned including Chopin, Beethoven, Victor Herbert and many others. Because of their meticulous fairness, American newsreels are featured without objection in theatres throughout the world.

"It is not to be assumed," continues Mr. Hays, "from attendance records of recent months that the American public has become so 'movie minded' that producers can afford to lessen their efforts for quality production. No greater error could be made. . . . The industry cannot afford ever to discontinue the vast entertainment research inherent in the making of pathfinding pictures."

Preliminary announcements have been received of the American Psychological Association meetings to be held at the University of Minnesota, Minneapolis, Minn., from September 1–4, 1937. Abstracts must be in the hands of the Secretary, Donald G. Paterson, University of Minnesota, on or before May 21, 1937.

Because of the wide interest shown in an experiment in improving the reading of retarded pupils, carried out with a federal subsidy in Theodore Roosevelt High School, New York City, the supervisors of the project, Dr. Stella C. Center and Miss Gladys L. Persons, have written a full account of the methods used and the results obtained. The book, entitled "Teaching High School Students to Read," and published by D. Appleton-Century Company, is said to be one of the most dramatic educational books that has appeared recently. The story of the project, which involved 500 retarded readers and 30 teachers, covers two years' work. The outstanding conclusion reached is that even seriously retarded pupils are capable of some progress in reading and hence of improving their chances of grasping other subjects. Both in diagnostic testing and corrective work, the latest scientific apparatus was employed, particularly the ophthalm-o-graph for photographing eye movements and the metron-o-scope for developing the habit of rapid, rhythmical reading. Teachers who must deal with reading retardation without benefit of a school remedial reading program will be especially interested in the analyses of causes of retardation and in the typical case studies given.

The United States Office of Education has recently prepared and distributed a pamphlet under the direction of Dr. James Frederick Rogers, Consultant in Hygiene and Health Education Specialist, listing more than 200 questions which school officers and parents might ask themselves concerning safety and health conditions in their schools. In view of the recent disaster in the New London, Texas, rural school, every precaution to prevent the recurrence of such a tragedy should be taken in every school in the country. Single copies of this 29-page pamphlet will be sent to anyone upon request.

The American Documentation Institute has been incorporated on behalf of leading national scholarly, scientific and informational societies to develop and operate facilities that are expected to promote research and knowledge in various intellectual ways. The board of trustees consists of Dr. Robert C. Binkley, Western Reserve University; Dr. Solon J. Buck, National Archives; Watson Davis, Director, Science Service; Dr. James Thayer Gerould, Librarian, Princeton University Library; Dr. Ludwig Hektoen, Chairman, National Research Council. Such a national organization was foreseen as an outcome of Science Service's documentation activities when they were begun in July, 1935, implemented with grants from the Chemical Foundation and conducted with the cooperation of the U. S. Naval Medical School, the U. S. Department of Agriculture Library, the Bureau of Census, the Works Progress Administration, the Library of Congress and other agencies. Bibliofilm Service has been conducted by Science Service in cooperation with the Library of the U. S. Department of Agriculture as a service to research workers.

BOOK REVIEWS

BROOM, M. E. *Educational Statistics for Beginning Students.* American Book Company, 1936. 318 pages.

There are many textbooks in educational statistics available for the use of persons under the necessity of assembling and of interpreting educational data, and it would seem that another textbook in the field might have little to contribute. There seems to be a place, however, for a textbook written for persons who use statistics rather than for persons who are interested in the development of statistical method.

The author states in his preface that the volume is dominated by two main ideas: first, that statistics intended for beginners in educational research should be simplified as much as possible, and that no unnecessary formulas and techniques should be presented; and second, that such beginning students of statistics should be made to encounter something of the philosophy of research and something of the products of research involving the use of statistical methods. The author has successfully presented in simple fashion the processes and techniques necessary for the collection and treatment of data commonly available to classroom teachers and to administrators. The chapters cover the minimum essentials: tabulation and classification, averages, measures of variability, measures of relationship, miscellaneous statistical constants, and graphic methods. Part II deals with the use and interpretation of educational statistics, with chapters on statistics and the classroom teacher, the uses of certain statistical measures, and statistical methods used in educational measurement, including methods of evaluating gains due to instructional methods, and statistical methods used by school administrators. There is included a selected bibliography and a group of tables for use in statistical computation.

Mechanically the volume is attractive. It is well bound, and the page make-up is excellent. Tables, formulas, charts, and graphs are printed clearly. Uniform practice was not followed in certain items throughout the book; for example, on pages 18 and 19 numbers are given, as 25 and 5, while on page 13 number words are used, as ten and fifty. There are several typographical errors. On page 83, in the tabulation at the bottom of the page, in the 90-row and in the 100-column, the number of cases should be 13, not 11. On page 85, in the second and third lines from the bottom of the page, there are errors in the printing of Σy^2 and Σx^2. On page 125, in the right-hand column of the table third line from the bottom, the entry should be plus 16.3, not minus 16.3.

It is the impression of the reviewer that this book will be of value to school administrators and classroom teachers and to those preparing to administer schools or to teach.

R. C. PERRY,
State College, San Diego, Calif.

DR. SOPHIA M. ROBINSON. *Can Delinquency be Measured?* New York. Columbia University Press, 1936. Pp. xxvi + 277. $3.00.

The author of this book is in charge of the Division of Neighborhood Statistics of the Welfare Council of New York City. The four years study presented here was a joint project of the Department of Sociology of Columbia University and the Research Bureau of the Welfare Council.

Delinquency is not invariably brought to public attention and is composed of many types of behavior deviation; accordingly it is not well reflected by official figures. As a result of delinquency area technique, latterly so popular, is essentially invalid as an indication of the extent of juvenile delinquency because cultural factors prove strongly influential upon the reporting of cases to courts or even to other agencies. Dr. Robinson's study attempts to find a truer method of indicating juvenile delinquency and also the characteristics of delinquent children. At present about one per cent of the children of eligible age, or 200,000 come before juvenile courts yearly; these figures do not sum up the problem, however, for the relation between court cases, agency cases, cases not reported, and cases not even brought to adult attention has not yet been established. Agency cases are shown to comprise for New York a group as large as court cases although differently made up because of the operation of group mores and the variable provision of unofficial agencies handling problem children.

Males above the age of ten, especially negro or those of Italian parentage, are most likely to appear in court; females, Jews, and Protestants are more likely to appear on agency records. Court records are made statistically useless by the use of waste basket terms such as 'ungovernable behavior'; a serious charge does not necessarily mean appearance in court or a minor charge reference to an agency. Both numbers and type of culprits are affected by such adventitious factors as police drives, special laws, agency set-ups, environmental factors such as the presence of railroad tracks. Poverty is challenged as a decisive factor as large numbers of children are reared in poverty without becoming delinquent and many children from families of moderate income are found on probation. The fundamental relation of truancy is also questioned, for not all truants become delinquents although delinquents may have been truants (55% of a comparable Sing Sing group admit truancy). New York city at least does not substantiate the theory that delinquency is a phenomenon of

interstitial areas nor do area rates seem meaningful. There seems no justification for the use of highly technical mathematical procedures, particularly where the unit of measurement is so inadequately defined.

The author needs to establish the importance of delinquency which goes unmeasured, for adult crime is apparently linked, not with delinquency *per se*, but with juvenile court and agency report which affect the child's sense of 'belongness' especially where institutionalization follows. The importance other than in romantic introspection of respectable faculty members is doubtful of "the misdeeds which the rest of us, reared in middle class families, would acknowledge now that we know the danger of punishment is past." At least a study would seem to be in order. Other follow up studies will suggest themselves from a reading of these provocative pages. The book is valuable in its plea for a new methodology in dealing with delinquency, the abandonment of the easy generalizations and resounding catchwords and the too great reliance upon statistical crutches which meet the present approval of most of us.

JAMES HARGAN,
Sing Sing Prison.

PATERSON, DONALD G., AND DARLEY, JOHN G. *Men, Women and Jobs, A Study in Human Engineering.* Minneapolis, Univ. of Minn. Press, 1936, pp. v + 145.

This interesting and socially significant little book reviews the work of the Employment Stabliziation Research Institute at Minneapolis and summarizes the various monographs issued by the Institute in the past few years. The basic data consist of case histories, including work history, clinical findings, both medical and psychological, occupational tests, interviewer's ratings and diagnosis of causes of unemployment. The detailed description of techniques and tests is supplemented with well-chosen photographs.

The first project described is a comparison of early and late unemployed workers by the general technique of item analysis. It develops that personal rather than economic reasons accounted for the earliest unemployment. Physical and mental factors are both involved, and the latter are shown graphically by the profile method.

A number of special types are then studied by the case method. Casual workers have low test scores and a third of them cannot be fully self-supporting. Non-modal workers, that is, those who have not been in one occupation long enough to be classified, are characterized primarily by youth. Another young group comprises the threshold workers who have not been able to get a start in any field. Unemployables constitute an appreciable number of the sample and some of them are physically handicapped. Of the entire sample of 2,000 workers only 45% were adequately

adjusted to their former work. Occupational ability patterns in the standard battery of tests are worked out for different types of employed workers by way of contrast to the unemployed.

Problems of re-education and training are considered critically. Such a program is now necessary in a surprisingly small number of cases. Those who took the recommended training, however, were more successful in their work than those who took training which was not recommended. A rather significant study is made of commercial correspondence courses where psychological tests reveal many persons attempting such courses beyond their capacity. The authors wisely recommend more discrimination on the part of the correspondence schools in determining whether the client has aptitude for the proposed vocation. A distinction is made between diagnosis and ambition. The Institute's own commercial training course revealed a definite relation between progress in the course and aptitude as indicated by prior tests. Differences between the Institute's and the usual commercial courses are arrayed in parallel columns to the obvious detriment of the latter.

The discussion of education digresses into problems of educational philosophy and interesting comment upon the platitude about equal opportunity. The authors endorse not the opportunity for everyone to become a physician, lawyer, or engineer but rather equal opportunity "to receive that training which will best fit him to support himself occupationally and participate socially in the group to which he will eventually belong." This is a clear statement with which few applied psychologists will disagree. Along somewhat the same line the authors disclose the fallacy in studies which show that college graduates receive higher compensation. The fundamental factor of intelligence is causally related to both the other variables.

The organization of the Employment Center is described. In developing their occupational ability patterns the criterion consists of the difference between a given occupational group and the population in general, rather than degrees of proficiency on the job. While this criterion is much more coarse than that usually applied within an industrial organization, it is reasonably satisfactory for the present type of problem, especially where the Center makes only the first selection and the prospective employer makes the final selection from a limited group of persons recommended by the Center. Some sound recommendations are presented for future testing programs in such centers, such as using valid tests, having them as integral part of the staff organization, having well-qualified examiners and using tests for positive as well as negative guidance.

The chapter on research is one of the longest in the book. The first portion deals with immediately practical research problems, such as ascertaining by means of trade tests how many alleged skilled tradesmen are

not even journeymen or discovering discrepancies between interviewers in different centers or finding comparatively small age differences in the tests. Basic research behind the scenes is also discussed. For example, norms in the various tests are secured in the "standard sample" which comprises individuals from the different vocations in the same proportion in which those vocations exist in the community. A classification for commercial and professional workers is developed, including social service, business d-tail, business contact, business service, professional or technological, ideational or lingual. One of the personality inventories (Bernreuter) is investigated and found ineffective in the type of problem met by the Institute. The same is true regarding Hepner's Vocational Interest Quotient, whereas Manson's and Strongs' Interest Schedules prove more satisfactory. Even though patterns have not been worked out for all the vocations the results of the factor analysis technique, proved valuable in interpreting patterns for various types of workers.

The book concludes with a plea for human engineering, of the type described, in school, industry, community and federal and state organizations. "Society is entitled to just as much professional skill in the handling of its personnel problems as in its health problems."

The book is concisely and clearly written. The topics are usually introduced by typical case studies followed by a more technical consideration. However, the general level of discussion is adapted to the layman. Except in the chapter on research there is little in the way of intricate techniques. Implicit in the work is a good deal of educational and social philosophy which should challenge the socially minded. Those who have not read the original monographs from the Institute or who lack the technical background to appreciate them will find this summary very profitable. It deals with a significant psychological and social project.

HAROLD E. BURTT,
Ohio State University.

NEW BOOKS AND PAMPHLETS RECEIVED

Books for review should be sent to James P. Porter, *Editor*, JOURNAL OF APPLIED PSYCHOLOGY, Ohio University, Athens, Ohio.

Compensating Industrial Effort. Z. CLARK DICKINSON. Ronald Press Company, New York, 1937. $4.50. 479 pp.

Competition and Cooperation. MARK A. MAY AND LEONARD W. DOOB. Bulletin No. 25, Social Science Research Council, 230 Park Avenue, New York, 1937. 191 pp.

Memorandum on Research in Competition and Cooperation. GORDON ALLPORT, GARDNER MURPHY AND MARK A. MAY. Social Science Research Council, 230 Park Avenue, New York, 1937.

Modern Discoveries in Medical Psychology. CLIFFORD ALLEN. Macmillan Company, New York, 1937. 280 pp.

Objective Analyses of Musical Performance. CARL E. SEASHORE, Editor. University of Iowa Studies in the Psychology of Music, Vol. IV. University of Iowa, February, 1937. 376 pp.

Personality and the Cultural Pattern. JAMES S. PLANT. The Commonwealth Fund, New York, 1937. $2.50. 432 pp.

School Adjustment of High School Pupils. LESTER A. KIRKENDALL. Bureau of Publications, Teachers College, Columbia University, New York, 1937. $1.60. 90 pp.

Selecting an Occupation. C. A. PROSSER AND R. H. PALMER. Information Book, 40 cents, 157 pp; Practice Book, 20 cents, 72 pp. McKnight & McKnight, Bloomington, Ill.

Studies in Psychology of Reading, Vol. I. JOSEPH TIFFIN. University of Iowa Studies in Psychology. Psychological Review Company, Providence, R. I., 1937. 149 pp.

Tohoku Psychologica Folia. T. CHIBA. Tohoku Imperialis Universitas, Sendai, Japan, 1936. 50 pp.

Directory of American Psychological Periodicals

(Continued)

Comparative Psychology Monographs—Baltimore: Johns Hopkins University Press. Subscription $5.00. 400 pages per volume. Knight Dunlap, Managing Editor. Published without fixed dates, each number a single research. Founded 1923.

Genetic Psychology Monographs—Provincetown, Mass.: The Journal Press. Subscription $7.00 per volume. One volume per year. Edited by Carl Murchison and an international cooperating board. Quarterly. Each number one complete research. Child behavior, animal behavior and comparative psychology. Founded 1925.

Psychological Abstracts—Eno Hall, Princeton, N. J.: American Psychological Association. Subscription $6.00. 700 pages ann. Edited by W. S. Hunter and R. R. Willoughby. Monthly. Abstracts of psychological literature. Founded 1927.

Journal of Genetic Psychology—Provincetown, Mass.: The Journal Press. Subscription $14.00 per annum; $7.00 per volume. 1,000 pages annually (2 volumes). Edited by Carl Murchison and an international cooperating board. Quarterly. Child behavior, animal behavior, and comparative psychology. Founded 1891.

Journal of Social Psychology—Provincetown, Mass.: The Journal Press. Subscription $7.00. 500–600 pages annually. Edited by John Dewey, Carl Murchison, and an international cooperating board. Quarterly. Political, racial, and differential psychology. Founded 1929.

Personnel Journal—Baltimore: The Williams & Wilkins Co. Subscription $5.00. 500 pages, approximately. Edited by Walter V. Bingham. Bi-monthly.

Mental Measurement Monographs—Baltimore: The Williams & Wilkins Company. Subscription $5.00. 400 pages annually. Edited by Knight Dunlap and Buford Johnson. Published without fixed dates, each number a single research. Founded 1925.

Studies in Psychology and Psychiatry—Baltimore: The Williams & Wilkins Company. Subscription $5.00. 400 pages annually. Edited by Edward A. Fass. Published without fixed dates, each number a single research. Founded 1926.

Journal of Educational Psychology—Baltimore, Md.: Warwick & York. Subscription $6.00. Edited by Jack W. Dunlap, Harold E. Jones, and Percival M. Symonds. Published monthly except June to August. Founded 1910.

Child Development—Baltimore, Md.: Williams & Wilkins Company. Subscription $5.00. 400 pages annually. Edited by Buford Johnson. Quarterly. Founded 1930.

Character and Personality—Durham, N. C.: Duke University Press. Subscription $3.00 per annum. Quarterly. Edited by Robert Saudek (London) and Ernest Seeman (Durham, N. C.). Founded 1932.

The Psychological Exchange—11 East 125th Street, New York, N. Y. Subscription $3.00 per annum. Bi-monthly. Edited by James Fargan. Founded 1932.

Journal of Psychology—Provincetown, Mass.: The Journal Press. Subscription $14.00 per annum; $7.00 per volume. 1000–1200 pages annually. Immediate publication for any accepted manuscript. Edited by Carl Murchison and an editorial board. Quarterly. The general field of psychology. Founded 1935.